Maximizing Spatial Reuse in Indoor Environments

Xi Liu

CMU-CS-11-103

December 2010

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Srinivasan Seshan, Co-Chair
Peter Steenkiste, Co-Chair
David Andersen
Konstantina Papagiannaki

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

CMU-CS-11-103

This research was sponsored by the National Science Foundation under grant numbers CNS-0721857 and CNS-0520192, IBM under grant number 5002518338, the U.S. Army Research Office under grant number DAAD190210389, and Intel. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of any sponsoring institution, the U.S. government or any other entity.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
DEC 2010		**00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Maximizing Spatial Reuse in Indoor Environments	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Carnegie MellonUniversity ,School of Computer Science,Pittsburgh,PA,15213**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Wireless technologies have gained tremendous popularity over the last couple of years. Such popularity causes dense and usually chaotic, i.e., unplanned and unmanaged deployments of wireless devices in indoor environments such as homes and offices. Due to the broadcast nature of wireless communication, wireless interference is becoming the most serious issue in these dense and usually chaotic wireless deployments. Existing techniques, however, cannot solve the interference problem at this scale and meet this growing demand for wireless capacity. In this dissertation, we propose to tackle the interference problem by optimizing spatial reuse (increase the number of simultaneous transmissions in an area without using additional frequencies) using directional transmission and power control. We propose two directional systems: DIRC for enterprise wireless networks with directional APs and omni-directional clients, and Speed for future wireless networks with directional APs and directional clients. While directional antennas can be very useful, there are certain cases where deploying directional antennas can be very difficult e.g., due to cost or size issues. Thus, in order to extend the optimization of spatial reuse to existing omni-directional antenna networks, we also propose Opera that uses the technique of power control to achieve spatial reuse for these networks. Both techniques of directional transmission and power control can allow multiple transmissions to happen simultaneously even in the same channel and even when the nodes are closely located in an area. While there are many systems use directional transmission and/or power control in outdoor application scenarios, achieving spatial reuse in our application scenario i.e., chaotic deployment of wireless devices in indoor environments, is decidedly more complicated. This is primarily due to the key characteristics of these scenarios Due to both rich scattering and unplanned node locations, the signal and interference patterns from the senders to the receivers are rather unpredictable. The problem is further complicated by the fact that today?s carrier sensing based medium access control (MAC) protocols interact poorly with both techniques. In this dissertation, we first show that contrary to conventional wisdom, directional transmission and power control can be very effective even in rich-scattered indoor environments. Then, we show how to build the three systems in a practical and lightweight fashion to exploit the capabilities of directional transmission and power control. Specifically, we use the SINR model to facilitate choosing the appropriate antenna orientations and power levels, and we also use a timeslot and timeslot

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**159**	

Keywords: wireless networks, directional antennas, power control, spatial reuse, wireless capacity, wireless performance, 802.11

For my parents and my dear wife Wenjun :-)

Abstract

Wireless technologies have gained tremendous popularity over the last couple of years. Such popularity causes dense and usually chaotic, i.e., unplanned and unmanaged, deployments of wireless devices in indoor environments such as homes and offices. Due to the broadcast nature of wireless communication, wireless interference is becoming the most serious issue in these dense and usually chaotic wireless deployments. Existing techniques, however, cannot solve the interference problem at this scale and meet this growing demand for wireless capacity.

In this dissertation, we propose to tackle the interference problem by optimizing spatial reuse (increase the number of simultaneous transmissions in an area without using additional frequencies) using *directional transmission* and *power control*. We propose two directional systems: DIRC for enterprise wireless networks with directional APs and omni-directional clients, and Speed for future wireless networks with directional APs and directional clients. While directional antennas can be very useful, there are certain cases where deploying directional antennas can be very difficult, e.g., due to cost or size issues. Thus, in order to extend the optimization of spatial reuse to existing omni-directional antenna networks, we also propose Opera that uses the technique of *power control* to achieve spatial reuse for these networks. Both techniques of directional transmission and power control can allow multiple transmissions to happen simultaneously even in the same channel and even when the nodes are closely located in an area.

While there are many systems use directional transmission and/or power control in outdoor application scenarios, achieving spatial reuse in our application scenario, i.e., chaotic deployment of wireless devices in indoor environments, is decidedly more complicated. This is primarily due to the key characteristics of these scenarios: Due to both rich scattering and unplanned node locations, the signal and interference patterns from the senders to the receivers are rather unpredictable. The problem is further complicated by the fact that today's carrier sensing based medium access control (MAC) protocols interact poorly with both techniques.

In this dissertation, we first show that contrary to conventional wisdom, directional transmission and power control can be very effective even in rich-scattered indoor environments. Then, we show how to build the three systems in a practical and lightweight fashion to exploit the capabilities of directional transmission and power control. Specifically, we use the SINR model to facilitate choosing the appropriate antenna orientations and power levels, and we also use a timeslot and timeslot reservation based MAC protocol (both centralized and distributed versions) to coordinate the APs and the clients. Finally, we evaluate our systems in several indoor testbeds to illustrate their effectiveness in practice.

Our contributions demonstrate that there exist practical and lightweight solutions to maximize indoor wireless spatial reuse.

Acknowledgments

First, I own my deepest gratitude to my advisors Srinivasan Seshan and Peter Steenkiste, who taught me how to conduct useful system research and how to present my work. I have been amazed by Srini's attention to details and thoughtful questions and insights. I have also been amazed by the fact that whenever Peter and I have some argument, it always turns out that Peter is correct. I hope I have inherited similar intuition during the last couple of years.

Second, I thank my other committee members and my collaborators that significantly contributed to the high level and technical details of my work: Dina Papagiannaki, Anmol Sheth, Michael Kaminsky, and Dave Andersen. I also would like to thank Hui Zhang, who gave me lots of "philosophical" advice on research and lifestyle during this time.

Third, I thank my fellow students and colleagues in CMU: Mukesh Agrawal, Eric Anderson, Fahad Dogar, Bin Fan, Bin Fu, Wenjie Fu, Leon Gu, Fan Guo, Dongsu Han, Favonia Hou, Glenn Judd, Kaushik Lakshminarayanan, Lei Li, Nan Li, Liu Liu, Amy Lu, George Nychis, Jeff Pang, Amar Phanishayee, Ning Qu, Kai Ren, Vyas Sekar, Runting Shi, Yanxin Shi, Kyung-Ah Sohn, Lin Xiao, Vijay Vasudevan, Xiaohui Wang, Yi Wu, Guangyu Xia, Hong Yan, Hormoz Zarnani, Xin Zhang, Yi Zhou, Jun Zhu. You guys made my PhD years more fun.

Finally, I thank my wife who comes all the way around the globe to be with me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Wireless technologies have gained tremendous popularity over the last couple of years. Such popularity causes dense and usually chaotic, i.e., unplanned and unmanaged, deployments of wireless devices in indoor environments such as homes and offices. At the same time, wireless applications that consist of streaming high definition multimedia applications, mobile displays, and wireless hubs require multiple Gbps of data bandwidth. Table 1.1 illustrates this growing demand for wireless capacity where density indicates the number of devices within interference ranges of each other. Examples include:

- Homes: It has been envisioned that in the future, every gadget and every appliance in your home will have a wireless device on it, which may add up to even one thousand radios in your future home. Typical applications may consist of heavy usages such as video streaming, wireless displays, wireless data storage, wireless hubs, and other light usages.
- Conference rooms: It is common to have hundreds of wireless devices in a conference room, consisting of laptops, netbooks, tablets, and smartphones. The size and density of the network makes it very difficult to manage these networks, and the performance is usually very poor. For example, in [79], the authors observed that more than 70% of the frames transmitted in a conference had at least one contender, even though the majority of the users only used wireless networks lightly.
- Metropolitan: It is quite common nowadays to be able to see tens of access points (APs) when you turn on your laptop and do a wireless scanning. In fact, [11] has shown that in several metropolitan areas, the density can be as high as 100s of APs within communication ranges of each other. With the introduction of new technologies such as femtocells, the density is expected to further increase.

Due to the broadcast nature of wireless communication and the fact that frequency is a relatively scarce resource, wireless interference is becoming the most serious issue in these dense and usually chaotic wireless deployments of hundreds and even thousands of devices. Existing techniques, however, cannot solve the interference problem at this scale and meet this growing demand for wireless capacity.

Table 1.2 illustrates an important trend over the last couple of years, which we believe is the key to tackle the interference problem: the availability of commercial and affordable multiple antenna wireless systems, namely directional antennas and multiple-input-multiple output (MIMO)

Scenario	# devices	Density	Planning
Home	1000 (projected)	100s (projected)	unplanned
Conference	100s	100s	planned with restriction
Enterprise	100s	10s	planned
Hotspots	100s	100s	unplanned across carriers

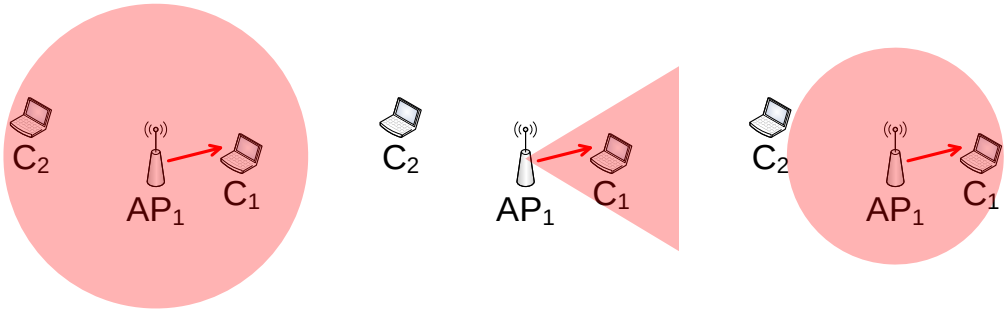
Table 1.1: Growing demand for wireless capacity in typical scenarios

Year	Antennas	Technology	Freq.	Rate	Directionality	Usage
1999	2	802.11	2.4GHz	2Mbps	omni	diversity
2003	2	802.11abg	2.4, 5GHz	54Mbps	omni	diversity
2007	3	802.11n	2.4, 5GHz	300Mbps	120°	MIMO
2007	8	802.11abg	2.4, 5GHz	54Mbps	45°	phased array
2010	40-60	WirelessHD,WiGig	60GHz	4Gpbs	5-9°	phased array

Table 1.2: Trend of existing and emerging wireless technologies

radios. While wireless devices that are more than two years old have only one or two antennas and use the second one for antenna diversity, i.e., use whichever antenna that is receiving the signal more strongly, many new devices have started to deploy the multiple antenna techniques of MIMO and directional antennas. We believe this trend will continue in several years and wireless devices will deploy more antenna elements, and is capable of performing directional transmission or MIMO beamforming.

Given these two trends, we propose to tackle the interference problem by optimizing spatial reuse (increase the number of simultaneous transmissions in an area without using additional frequencies) using *directional transmission* or, in general, beamforming. In this dissertation, we propose two directional antenna systems: DIRC for enterprise wireless networks with directional APs and omni-directional clients [75], and Speed for future wireless networks with directional APs and directional clients [76]. While directional antennas can be very useful, there are certain cases where deploying directional antennas can be very difficult and omni-directional antennas are more appropriate, e.g., due to cost or size issues. Also, we would like to extend the optimization of spatial reuse to existing omni-directional antenna networks. Thus, for omni-directional antenna systems, we also propose Opera. Opera uses the technique of *power control* to achieve spatial reuse for these networks [73]. Both techniques of directional transmission and power control can allow multiple transmissions to happen simultaneously in the same channel and even when the nodes are closely located in an area. Figure 1.1 shows an example with a sender AP_1, its intended receiver C_1, and an unintended receiver C_2. Figure 1.1(a) shows that while AP_1 is transmitting to C_1, it will introduce strong interference at C_2 that prevents C_2 from receiving any frames. As shown in Figure 1.1(b)&(c), directional transmission can confine AP_1's communication and interference range to a narrow region and power control can shrink the circular range of



(a) No spatial reuse (b) Directional transmission (c) Power control

Figure 1.1: Example of achieving spatial reuse through directional transmission and power control

AP_1's communication. By deploying either technique, AP_1 does not interfere with C_2 anymore, thus enabling spatial reuse in this example scenario.

While many systems use directional transmission and/or power control to achieve various goals in outdoor application scenarios, achieving spatial reuse in our application scenario, i.e., chaotic deployment of wireless devices in indoor environments, is decidedly more complicated. This is primarily due to the key characteristics of these scenarios: Due to both rich scattering and unplanned node locations, the signal and interference patterns from the senders to the receivers are unpredictable. In Chapter 2, we show examples (Figure 2.2) of two typical receive patterns for directional antennas which illustrates this unpredictability. This unpredictability significantly complicates the choices of the right antenna orientations and power levels. In fact, due to this reason, the conventional wisdom has been that directional antennas are not effective in indoor environments at all. The problem is further complicated by the fact that carrier sensing based medium access control (or MAC) protocols, that are widely used in today's wireless networks, interact poorly with both techniques. Thus a new MAC protocol needs to be designed to exploit the spatial reuse opportunities provided by both techniques.

1.1 Current Approaches

From a physics perspective, there are three orthogonal approaches to avoid interference: time, frequency, and space. Carrier sensing, i.e., sense the carrier before transmit a frame, is one example solution in the time domain. These solutions, however, do not improve spatial reuse. The basic idea of the frequency domain is that devices operate in orthogonal frequencies do not interfere with each other. However, frequency is still a relatively scarce resource, e.g., Wi-Fi only has access to three orthogonal channels in 2.4GHz and eleven orthogonal channels in 5GHz. Recently, the FCC regulations open up the unused UHF spectrum adding up to 180MHz, referred to as "White Spaces". Though the interference problem can be greatly alleviated with the availability of the new spectrum, we believe that these solutions alone will not suffice to deal with the interference problem at the scale of thousands of radios, especially due to the signal propagation properties of the UHF spectrum. Also note that research in this area is still in a

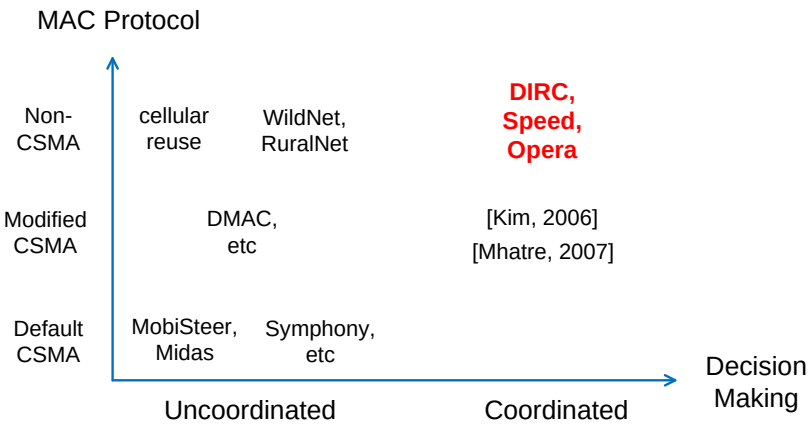


Figure 1.2: Categorization of existing approaches on directional transmission and power control

very early stage where interactions or interference among the APs are largely ignored. It is our belief that the ultimate solution to the interference problem at the scale of thousands of radios will combine ideas from all three domains. Our work in this dissertation contributes to the space domain, i.e., maximizing spatial reuse at no cost of extra frequency bands.

The idea of improving spatial reuse using directional transmission and/or power control is not new. For example, cellular networks deploy both techniques. Many other wireless systems that use either directional transmission or power control have also been proposed. In order to distinguish our contributions from prior efforts, we categorize existing systems and our proposed systems (DIRC [75], Speed [76], and Opera [73]) along two dimensions:

1. Whether the decision of antenna orientations or power levels is determined in an uncoordinated way to optimize individual link metrics, or in a coordinated way to optimize network-wide goals.
2. What the MAC protocol is, i.e., default carrier sensing multiple access (CSMA), CSMA with some form of modification or adaptation, or non-CSMA protocols.

We only provide a high level overview of different systems with different solutions in this chapter, and we will compare different solutions in more details in the next chapter.

Figure 1.2 shows this categorization for all existing systems that optimize spatial reuse and our proposed solutions (DIRC, Speed, and Opera marked in bold and red). For directional antenna systems, prior efforts (cellular networks, WildNet, DMAC, MobiSteer, Midas, etc. in Figure 1.2) focus on local optimizations or uncoordinated decisions where each sender and/or receiver use the antenna orientations that maximize the signal strength at the receiver. For power control systems, most existing systems (Symphony, etc. in Figure 1.2) decide to minimize the power level on each sender such that the signal strength at the receiver is just enough to receive the frames from that sender. Other power control systems make coordinated decisions to optimize network-wide goals and rely on CSMA protocols with some form of modification or adaptation, i.e., tuning the carrier sensing thresholds or clear channel assessment (CCA tuning). To distinguish our contributions, all our three systems optimize network-wide goals through node coordination and rely on non-CSMA MAC protocols. We will explain why our solutions (coordinated solutions with non-CSMA protocols) achieve better performance (next section) and

we quantify the gain in practice through measurement-based studies (in later chapters). We now discuss the two dimensions in Figure 1.2 in more details.

1.1.1 Uncoordinated vs. Coordinated Decisions

The first dimension is the choice of antenna orientations and/or power levels. The most straightforward way to choose the antenna orientation is to maximize the signal strength at the receiver; and the most straightforward way to choose the power level is to minimize the power level such that the signal strength at the receiver is just enough to receive from the sender. Both these approaches are uncoordinated solutions that optimize local link metrics. Note that the application scenario we target is chaotic deployment of wireless nodes in indoor environments. Here are several reasons why existing systems make uncoordinated decisions, and why they cannot be applied to our application scenario:

- Outdoor directional systems (WildNet [90], DMAC [65]): For these systems, there is usually only one antenna orientation available, i.e., along the line-of-sight (LOS) path, thus antenna orientation is relatively simple. In indoor environments, however, there usually exist multiple paths between the senders and the receivers.
- Noise dominated indoor directional systems (Midas [12], Xirrus sectored Wi-Fi array [8], Phocus array [3], Ruckus BeamFlex [5]): If noise dominates interference in the network, e.g., sparse wireless networks with very light traffic, optimizing signal strength at the receivers will optimize link throughputs and network capacity as well. However, for dense deployments of wireless nodes, the networks become interference dominated, and managing interference is the key to improve network capacity. Existing commercial products (Xirrus sectored Wi-Fi array, Phocus array, and Ruckus BeamFlex) primarily focus on noise dominated application scenarios because today's wireless networks are still mostly noise dominated. We believe that given the popularity of wireless devices, more wireless networks will become interference dominated. Also note that spatial reuse will be improved even if these directional antenna make uncoordinated decisions. However, as we will show in Chapter 2, they fail to maximize spatial reuse in these indoor environments without coordination.
- Power controlled systems to save energy (Symphony [95]): If the primary optimization goal is to save energy, minimizing power levels is the appropriate solution and spatial reuse is generally only a side effect of these systems. However, our goal is to optimize spatial reuse. Thus these solutions will not suffice.
- Systems in licensed spectrum (cellular frequency reuse): The spectrum used in cellular networks is licensed. This means that the network operators have the full control over the locations of the cellular towers, which are carefully planned to form hexagons, so interference within the network is predictable. However, in chaotic network deployments, even the placement of APs is unplanned. Even for enterprise networks where the locations of the APs can be carefully planned, the rich scattering of the indoor environments makes the receive patterns unpredictable.

The key difference between our application scenario and that of earlier work is that the signal and interference patterns at the receivers are unpredictable. Due to the unpredictability, uncoor-

dinated decisions that ignore the interactions among the senders cannot maximize spatial reuse in these scenarios, and coordinated optimizations are necessary.

One related issue is centralized vs. decentralized solutions to choosing the antenna orientations and/or power levels. Centralized solutions can better support coordinated decisions, but can only be applied to certain application scenarios such as enterprise networks. On the other hand, distributed solutions can be applied to much broader application scenarios, but coordination becomes much more difficult. In DIRC, directional APs are connected through a wired network and are centrally controlled by a controller. Both Speed and Opera are fully distributed systems.

1.1.2 Medium Access Control (MAC) Protocol

The second dimension is the MAC protocol. Even when the most appropriate antenna orientations and power levels are chosen to enable spatial reuse at the physical layer, a properly designed MAC protocol is further needed to exploit these opportunities. The most important requirement for the MAC protocol is the ability to identify the active transmissions.

Unfortunately, while CSMA-based MAC protocols work reasonably well in 802.11 omni-directional antenna networks, they interact poorly with both directional transmission and power control. For directional transmission, the key assumption made in the carrier sensing based mechanisms, i.e., the carrier sensed at the sender is a good indicator of the interference level at the receiver, fails in directional antenna networks. For power control, tuning power levels may cause link asymmetry, i.e., one sender can carrier sense another sender but not the other way round. Joint tuning of CCA and power levels can alleviate this problem but is still undesirable because it cannot satisfy the requirement of identifying the active transmissions. Furthermore, it significantly complicates the problem, and has many implementation limitations. In summary, CSMA-based MAC protocols do a poor job in identifying the active transmissions.

Non-CSMA protocols have also been adopted in earlier systems for various reasons. For example, cellular networks use various protocols such as Global System for Mobile Communications (GSM) and Code Division Multiple Access (CDMA) that are optimized for voice traffic, i.e., continuous traffic with a fixed bandwidth requirement. WildNet [90] and RuralNet [97] use simplified MAC protocols that largely ignore link interactions. This is acceptable because of the very high isolation from the narrow beam parabolic antennas in these systems. Clearly these solutions cannot identify active transmissions and cannot be applied to our systems.

It is worth noting that many systems that modify the default CSMA behavior (such as [65] which introduces a new directional MAC protocol based on virtual carrier sensing and [81] which tunes CCA thresholds) cannot fairly share the channel with 802.11 systems. Earlier work that simply uses the default CSMA protocol can sometimes be problematic as well (e.g., link asymmetry caused by power control). Interestingly, even though our systems rely on non-CSMA MAC protocols, they can in fact co-exist and fairly share the channel with 802.11 systems, as explained in Chapter 5.

1.1.3 Other Dimensions

While the two dimensions of coordination and MAC protocol is sufficient to distinguish our work from prior work, there are two other major dimensions as well to categorize existing solutions

and our proposed systems.

Optimization Goals

As mentioned earlier, prior efforts have various optimization goals such as increasing communication range [90], minimizing average latency [81], and saving energy [95]. The primary goal of our work is to optimize spatial reuse.

Optimizing spatial reuse, in fact, can increase the link throughputs on all transmissions in the system. This is in contrast with work that trade fairness for higher network capacity [21, 93, 110]. Also, whether these improvements on all links will be perceived as "better" or "worse" fairness (as the improvements for some transmissions can be much higher than for others because the sender and the receiver are in better locations) depends on the definition of fairness. We will discuss two definitions in Chapter 2.

Because we optimize spatial reuse with node coordination, our systems mainly target nomadic usage, i.e., clients that may appear at multiple locations across the network but are used in stationary positions (laptops). In the evaluation of the systems, we find that the performance of our systems does not degrade too much (less than 30% on average) in events with low degree of mobility, e.g., users walk around. However, our systems are not designed to handle highly mobile scenarios, e.g., vehicular ad-hoc networks (VANETs) [12, 87]. We will discuss how can the systems better handle mobility in Chapter 8.

Interference Model

While many earlier systems rely on the simple range based model [95], our systems rely on a model that is more accurate for today's hardware: the signal to interference plus noise ratio (SINR) model. The description and comparison of both models are presented in Chapter 3.

A related issue is the path loss model. While several systems rely on free space path loss model [38], we make no assumption on the environment, the location of the APs and the clients, the received signal strength patterns, or path losses between senders and receivers. All signal strength readings are directly measured in our systems.

1.2 Thesis and Approach

In this dissertation, we show the validity of the following thesis:

> *There exist practical and lightweight designs for both directional transmission and power control in which nodes coordinate with each other to effectively optimize network-wide spatial reuse in chaotic and dense wireless networks even in rich scattered indoor environments.*

In the previous section, we have shown that our systems differ from prior efforts along two dimensions: nodes in our systems make coordinated decisions to optimize the network-wide goal of spatial reuse, and our systems use non-CSMA MAC protocols. Designing systems in this

point of the two-dimensional design space, however, is very challenging. In fact, there are two primary challenges in system design, one for each dimension.

The first challenge is how to make coordinated decisions and, more importantly, how to make these decisions quickly. The naive way to optimize network-wide spatial reuse or capacity is to do an exhaustive search that micro-benchmarks the whole space of antenna orientations and/or power levels. Unfortunately, this exhaustive search is simply too slow and impractical for directional antenna systems and for power controlled systems. Our approach is to measure the signal strength from the APs to the clients, and then use the SINR model to predict the link throughputs and the network capacity. Using this approach, the number of measurements necessary to make the coordinated decisions can be significantly reduced, making this process quick enough to be practical.

The second challenge is how to design a non-CSMA protocol to effectively exploit the capabilities of directional transmission and power control. The requirements for the MAC protocol are two-fold:

- to provide the coordination mechanism for the nodes to choose the right antenna orientations and/or power levels, and
- to allow non-interfering transmissions to happen simultaneously and to avoid interfering transmissions from occurring at the same time.

Note that since the choice of the antenna orientations and/or power levels in the first requirement will affect the interference conditions in the second requirement, these two requirements need to be considered together. Our approach is to use a MAC protocol that is based on timeslots and timeslot reservations. In this MAC protocol, we use the information collected during the signal strength measurements to ensure that the protocol allows as many transmissions as possible, and that no interference happens among the transmissions. DIRC uses the centralized version of this MAC protocol, and Speed and Opera use the distributed version.

One related challenge of designing a non-CSMA MAC protocol, regarding its deployability, is to allow nodes running this non-CSMA MAC protocol to co-exist and fairly share the channel with nodes running CSMA-based protocols. Since CSMA-based protocols are the most popular ones in unlicensed frequency bands, any newly designed non-CSMA MAC protocol that monopolize the channel is unlikely to be deployed. We address this challenge by implementing the non-CSMA MAC protocol using carrier sensing mechanism, i.e., using carrier sensing to implement non-CSMA functions, as we will show in Chapter 5.3.5. By disabling random backoff, we show that we can implement the non-CSMA MAC protocol in a way that it can co-exist and fairly share the channel with other CSMA-based protocols.

1.3 Our Contributions

In addition to showing the validity of the thesis, we also make the following contributions:

- We show that both directional transmission and power control can be very effective, i.e., significantly improve spatial reuse over naive solutions, even in rich scattered indoor environments if the systems are carefully designed.
- Using the SINR model, we greatly reduce the number of measurements to choose the right

antenna orientations and power levels in our systems to optimize spatial reuse in a quick, lightweight, and practical fashion.

- We design and build a timeslot and timeslot reservation based MAC protocols (both centralized and decentralized) to exploit the spatial reuse opportunities, i.e., to prevent (directional) hidden terminal and (directional) exposed terminal problems, based on the observations that CSMA-based protocols interact poorly with both directional transmission and power control.
- Using a simple SINR based metric, the separation metric (including distance and angular separation), we can quickly evaluate the effectiveness of directional transmission and power control in any particular environment, and can guide the placement of directional APs in that environment.

1.3.1 Thesis Scope

It is worth noting that the contributions in this dissertation have only been experimentally validated in specific environments: saturated traffic in dense wireless network deployments in 2.4GHz frequency bands. In Chapter 7.1 and Chapter 7.2, we identify three major factors that affect the effectiveness of directional transmission and power control and discuss the applicability of our work to other environments. In summary, we believe that both directional transmission and power control will still be very useful in other environments, e.g., UHF and 60GHz bands. Further studies are needed to quantify the actual benefits.

1.4 Outline

The rest of the dissertation is organized as follows:

- In Chapter 2, we present measurement-based studies to motivate the three systems proposed in this dissertation. Specifically, we show that both directional transmission and power control can be effective in rich scattered indoor environments, and that naive solutions are unable to exploit these benefits.
- In Chapter 3, we present the SINR interference model, which is the key model in our systems to significantly reduce the number of measurements needed to make the coordinated decisions. We also discuss several limitations of the model, and how we modified the model to make it work well in practice.
- In Chapter 4, we present the design, implementation, and evaluation of DIRC, which is a directional antenna system designed for enterprise wireless networks where APs are directional and clients are omni-directional.
- In Chapter 5, we present the design, implementation, and evaluation of Speed, which is a directional antenna system designed for future wireless networks with directional APs and directional clients.
- In Chapter 6, we present the design, implementation, and evaluation of Opera, which is a power control system for existing omni-directional networks.

- In Chapter 7, we present three major factors that affect the effectiveness of directional transmission and power control in any particular environment, and propose separation metric to determine this effectiveness, and guide the placement of directional APs in any environment.
- In Chapter 8, we summarize the contributions and limitations of our systems, and outline several future directions.

Chapter 2

Characterizing Spatial Reuse in Indoor Environments

In the previous chapter, we have shown the categorization of existing systems (and our proposed systems) that optimize spatial reuse, and explain why prior efforts fail to maximize spatial reuse in our application scenario. In this chapter, we use measurement based studies to illustrate the effectiveness of directional antennas and power control and the ineffectiveness of naive solutions in several real indoor environments. Then, we motivate three systems, DIRC, Speed, and Opera, in the design space, based on the measurement studies. Finally, we discuss related work on measuring the performance of directional antennas and power control.

2.1 System Model and Notations

Before presenting the measurement based studies, we present the system model and notations used in this chapter and the rest of the dissertation.

System Model

In our systems, we assume an infrastructure wireless network (we discuss the possibility and implications of applying our techniques to mesh or adhoc networks in Chapter 8) where there are N APs $\{AP_1, AP_2, ..., AP_N\}$ and M clients $\{C_1, C_2, ..., C_M\}$ ($M \geq N$). Each AP has K_{AP} directions and each client has K_C directions. Note that for omni-directional APs, $K_{AP} = 1$, and for omni-directional clients, $K_C = 1$. We use $S(AP_i, C_j, K_{AP_i}, K_{C_j})$ to denote the received signal strength from AP_i to C_j with orientation K_{AP_i} on AP_i and K_{C_j} on C_j. The number of power levels on the APs is P_{AP}, and the power level used on AP_i is P_{AP_i}. Also we assume there are DR discrete data rates. Note that the notations may be slightly different in different contexts, which we will highlight in later chapters.

Co-channel Operation

We assume that all nodes (APs and clients) in our systems operate in the same channel. We ignore nodes operating on orthogonal channels as they do not interfere with each other. This also

indicates that each one of our experimental testbeds represents one of the multiple orthogonal channels in that particular environment. For example, the experimental results obtained from our testbeds will reflect that from larger networks featuring 3 times (in 2.4GHz) or 11 times (in 5GHz) more nodes.

Downlink vs. Uplink Traffic

Since the majority of the traffic in wireless networks is from APs to clients, i.e., downlink, we optimize our systems for downlink traffic. Uplink traffic in DIRC, i.e., from clients to APs, falls back to use omni-directional mode and 802.11 protocol. Uplink traffic in Speed and Opera can be part of the protocol, but we did not implement this in the systems. Also throughout the dissertation, we use AP and sender/transmitter/tx interchangeably, and client and receiver/rx interchangeably. Thus we also use the following notations for different directional antenna configurations: *dir-tx,dir-rx* for directional APs and clients, *dir-tx,omn-rx* for directional APs only, *omn-tx,dir-rx* for directional clients only, and *omn-tx,omn-rx* for omni-directional APs and clients.

Path Loss Model

The key property of the chaotic and dense indoor wireless networks is that the received signal strength at the clients are unpredictable, both because of the unplanned placement of both the APs and the clients and because of the rich scattering in indoors environments. This property is observed in our experiments and has been pointed out in previous work [89, 129] as well. Thus it is important to note that in this dissertation, we make no assumption on 1) path loss and/or received signal strength patterns, 2) or the location of the APs and the clients, 3) or the environment. All signal strength readings are directly measured in our systems.

Network Capacity and Fairness

While fairness has an agreed-upon definition in wired networks [37, 61, 111] or operating systems [71], there are conflicting notions of fairness in wireless networks. One fairness definition is Jain's fairness index [54], defined as:

$$fairness = \frac{(\sum x_i)^2}{n \sum x_i^2}$$

Jain's fairness index ranges from $1/n$ to 1, and it will be 1 if all transmissions receive the same throughput. This index is examining the "absolute" fairness, and is used in some wireless research [51]. Another fairness definition is max-min fairness [21, 93], defined as:

$$fairness = \min_i(\text{link thp}_i)$$

Basically, according to max-min fairness, the fairness of a scenario with starved links will be bad.

Generally, there is a tradeoff between network capacity and fairness. For example, consider the simple scenario with 9 transmissions, $t_1 - t_9$. In this network, only t_1 interferes with all

other transmissions, but otherwise all transmissions can occur simultaneously. In this case, to maximize network capcaity, t_1 should not be allowed transmitting at all, and all the other 8 transmissions should be allowed simultaneously, achieving a capacity of 8. On the other hand, to maximize fairness, we should allow t_1 to transmit half of the time and allow all other transmission to send the other half of the time, acheving a capacity of 4.5. From this example, we can see that achieving maximum capacity will cause starvation and achieving maximum fairness will greatly reduce network capacity. In fact, we believe neither approach is reasonable.

In our systems, we adopt the implicit fairness model in the 802.11 systems. In 802.11 systems, fairness is ensured in the sense that each node has equal opportunity to grab the channel if the channel measured at the sender is clear. In our systems, by enabling spatial reuse, we improve the link throughputs on all transmissions. Interestingly, according to Jain's fairness index, the fairness of our systems may be even lower than that in 802.11 systems (though everyone is getting higher throughputs), but the max-min fairness is improved.

2.2 Experimental Setup

The measurement studies presented in this section and the rest of the dissertation is carried out in various different physical locations with different number of APs and clients and different antenna configurations. Table 2.1 summarizes the experimental setup for all scenarios. The columns "AP ant." and "client ant." show the antenna setups we deploy on APs and clients ("all avail." indicates we use all setups including omni-directional). The map and details of each setup are shown in later chapters. The wireless cards used in the experiments use either Atheros 5212 or 5213 chipsets, and both chipsets have similar features. The antenna patterns of all the antennas is shown in Figure 2.1. Next we examine the capabilities of existing directional antenna and power control techniques.

2.2.1 Directional Antennas

Generally speaking, there are three types of directional antennas: parabolic antennas, phased array antennas, and patch antennas. Table 2.2 summarizes these three different types of directional antennas.

Parabolic antennas can usually form very narrow beams and are inexpensive, but they are very bulky in size and are mechanically steered. Due to their size, they are primarily used in outdoor applications to extend communication range [90, 97, 108]. In our measurement studies, we used a 16° parabolic antenna, and the picture and the antenna pattern are shown in Figure 2.1(a)&(b).

Phased array antennas can electronically steer their beams towards any direction, but they are usually very expensive and can only be deployed on APs (the size is too large on the clients). In our measurements and system design, we use the Phocus phased array antenna [3, 20, 26, 27, 68, 87, 96, 112, 118]. The Phocus array has 8 antenna elements and can form 45° beams. In fact, the users can specify the amplitude and the delay on each antenna element and thus could form irregular antenna patterns [68, 96, 118]. In our systems, however, we simply use the default factory loaded antenna patterns (we discuss how our systems may be extended to use irregular antenna patterns in Chapter 8). The picture of the antenna and its antenna patterns (both directional and

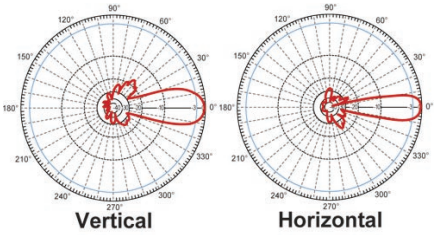
Scenario	# APs	# clients	AP ant.	client ant.	Map
Campus 1	3	6	45°,omni	omni	4.4
Lab 1	3	6	45°,omni	omni	4.4
Campus 2	6	6	16,35°,omni	$4\times35°$ fan beam, omni	5.8
Lab 2	6	6	16,35°,omni	$4\times35°$ fan beam, omni	5.8
Topo 1-5	2	2	35°,omni	$4\times35°$ fan beam, omni	5.8
Campus 3	6	6	all avail.	omni	5.8
Lab 3	13	10	all avail.	omni	5.8
Campus 4	20	28	60°,omni	omni	7.2
Outdoor	1	1	all avail.	omni	an outdoor field
Campus 5	9	5	omni	omni	N/A
Campus 6	5	7	omni	omni	N/A
City	11	9	omni	omni	N/A

Table 2.1: Experimental setup for all scenarios

Type	Size	Price	Steering	Usage
Parabolic	large	inexpensive	mechanical	outdoor only
Phased-array	medium	very expensive	electronic	APs
Patch	small	inexpensive	mechanical	APs and clients

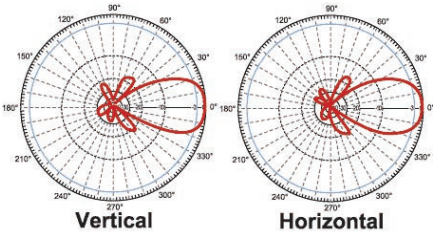
Table 2.2: Directional antenna technologies

(a) parabolic antenna (with turntable)

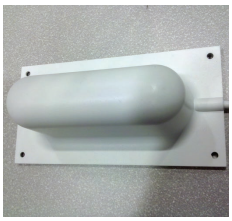


(b) pattern (16°)

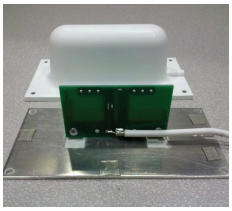
(c) panel antenna



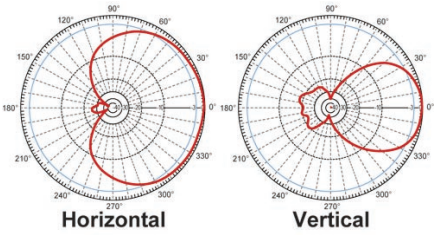
(d) pattern (35°)



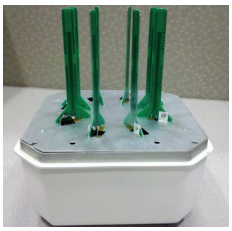
(e) fan beam antenna (outside)



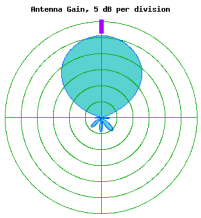
(f) fan beam antenna (inside)



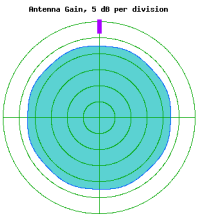
(g) pattern (35° in one plane and 135° in another)



(h) Phocus array

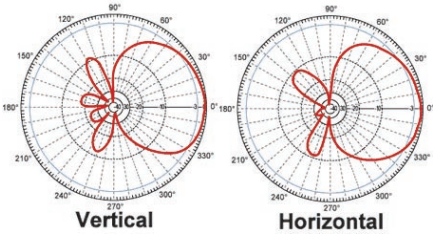


(i) pattern (45°)



(j) pattern (omni)

(k) patch antenna



(l) pattern (75°)

Figure 2.1: Directional antennas and antenna patterns

(a) RSS from -63 to -45dBm (b) RSS from -88 to -75dBm

Figure 2.2: Typical received antenna patterns in indoor environments (for 45° directional antennas)

omni-directional) are shown in Figure 2.1(h)-(j). The Phocus array antenna can electronically steer its antenna orientation in $100\mu s$, thus orientation can be changed frequently. Unfortunately, the antenna orientation is steered in the driver while when the frame gets transmitted is determined in the hardware abstraction layer (HAL) and thus accurate control of per-packet antenna steering is currently not supported.

Patch antennas are small and inexpensive antennas that form a fixed sector of directional beam. The small patch antennas can form 75° beams in both the horizontal and the vertical planes. We also use a larger patch antenna of 35° beamwidth and a type of fan beam patch antenna that forms 35° beam in vertical plane and 135° beam in horizontal plane. The pictures and the antenna patterns for these three types of antennas are shown in Figure 2.1(k)&(l), Figure 2.1(c)-(d), and Figure 2.1(e)-(g), respectively.

In our measurement based studies, we use a turntable to mechanically steer the parabolic and patch antennas to the desirable orientation. The turn table can mechanically orient itself to 64 different directions, but we only use 32 different directions (with 11.25° between adjacent directions). The picture of the turntable is shown in Figure 2.1(a).

Figure 2.2 shows typical received antenna patterns in an indoor environment with 45° directional senders and omni-directional receivers. In each figure, the center point shows the minimum RSS, the farthest point shows the maximum RSS, and both numbers are included in the caption of each figure. And the signal strength of any point within the range is shown proportionally (in terms of dBm) in each figure, i.e., a point half way between the maximum and minimum RSS indicates the RSS level of $(max+min)/2$. The resulting received antenna patterns, in contrast with the clean shape shown in Figure 2.1(i), illustrate that there are usually multiple paths between the sender and the receiver, and the number of paths and their locations are unpredictable.

2.2.2 Power Control Capability

The wireless cards with Atheros [2] chipsets can support per-packet power control, i.e., specify the power level in the frame descriptor and the hardware will use the specified power level when transmitting the frame.

A related capability is CCA tuning. Currently, only the AP version of the firmware and the driver for Intel cards (note that there are separation firmware and driver for AP and client side) support reasonably accurate CCA tuning capability (unfortunately the support for both the driver and the firmware has been discontinued). Even in this case, the CCA tuning granularity is coarse, i.e., per-packet tuning is not supported. Also, any implementation that tunes the CCA thresholds does not work very well with 802.11 nodes, as the nodes that have a high CCA threshold cannot carrier sense the 802.11 nodes and may monopolize the channel.

2.2.3 Methodology

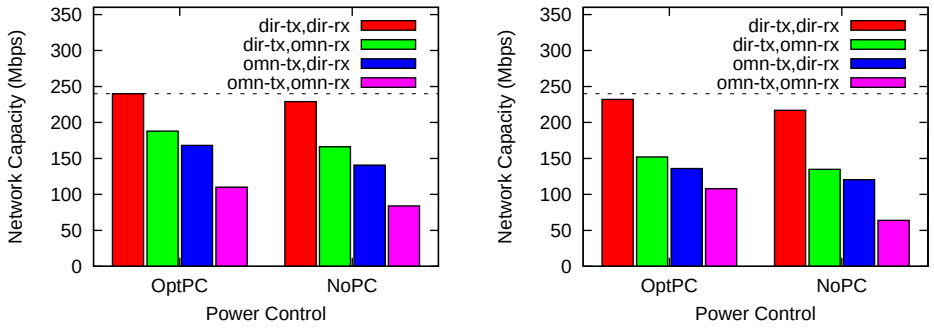
As mentioned before, we carried out our measurement studies in various scenarios as summarized in Table 2.1. In each scenario, we collect the signal strength measurements from all APs to all clients with all possible directions on APs and clients, i.e., the complete table of $S(AP_i, C_j, K_{AP_i}, K_{C_j})$. In addition to signal strength measurements, we also directly measure the network capacity in Campus 1 and Lab 1 scenarios, by counting the number of frames that can be received from the receivers with the senders active simultaneously. For the rest scenarios, the network capacity is computed using the SINR model based on the collected signal strength measurements. We rely on the measurements collected on the clients during this process, and do not rely on reciprocity. This process is described in more details in Chapter 3.

We first look at the problem of orienting the directional antennas. Directional antennas have primarily been used in outdoor deployments, where the LOS orientation of the antenna towards the receiver provides both the best performance and the best signal strength [90, 97]. In indoor environments, however, the LOS path may not exist because of obstructions between the sender and the receiver. And we evaluate two alternative orientation algorithms for indoor environments:

- MaxSNR ([75, 76]): The sender and/or receiver use the antenna orientations to maximize the signal strength at the receiver. This is the most straightforward solution that most existing indoor directional deployments use (Xirrus sectored WiFi array [8], Phocus array [3], and Ruckus BeamFlex [5]). Note that this is an uncoordinated algorithm that optimizes a single link metric.
- MaxCAP ([75, 76]): The senders and/or receivers use the antenna orientations to maximize the network capacity at all receivers. This involves an exhaustive search on the entire antenna orientation space, and the antenna orientations with the highest network capacity are chosen. Note that this is a coordinated algorithm that optimizes a network-wide goal, and can achieve the upper bound performance in that particular scenario.

Then for power control, we also examine three simple algorithms:

- NoPC: All senders use the same power levels.
- MinPC ([73]): The sender uses minimum power level just enough to reach the receiver with the highest data rate. This is the naive solution that most power control systems



(a) Network capacity (Campus 2)

(b) Network capacity (Lab 2)

Figure 2.3: Performance benefit of directional transmission and power control

deploy. Note that this is an uncoordinated algorithm that optimizes a single link metric.

- OptPC ([73]): The senders use the power levels that maximize the network capacity at all receivers. This involves an exhaustive search on the entire space of all possible power level on all senders, and the power levels with the highest network capacity are chosen. Note that this is a coordinated algorithm that optimizes a network-wide goal, and can achieve the upper bound performance in that particular scenario.

The previous algorithms are used to determine how to choose antenna orientations or power levels to enable physical layer spatial reuse. Also, in order to isolate the effects of the MAC layer from the spatial reuse opportunities enabled by the physical layer, we also examine the performance of different MAC layers:

- CSMA: This is the 802.11 MAC protocol, i.e., the default CSMA protocol without tuning CCA thresholds (use the default CCA thresholds).
- CSMA with CCA tuning: This is the CSMA protocol with CCA threshold tuning. For this protocol, we do an exhaustive search on all possible CCA thresholds to maximize capacity.
- OptMAC ([73, 75, 76]): This is the optimal scheduler. The optimal scheduler does an exhaustive search on all possible schedules (schedules of which senders to transmit, along with antenna orientations and/or power levels), and choose the schedule that has the highest spatial reuse. Note that this algorithm can achieve the upper bound MAC performance in that particular scenario.

2.3 Performance Benefits of Directional Transmission and Power Control

In this section, we present the performance benefits of directional transmission and power control in two scenarios. Also, we will to focus on the spatial reuse opportunities enabled in the physical layer. Thus we factor out the MAC layer effects by using OptMAC protocol on all antenna orientation and power control algorithms.

Figure 2.3 shows the network capacity (the sum of all link throughputs, and saturated down-

link only) on the six transmissions in Campus 2 and Lab 2 scenarios, with four different antenna configurations (*dir-tx,dir-rx*,*dir-tx,omn-rx*, *omn-tx,dir-rx*,*omn-tx,omn-rx*) and with and without power control. In both scenarios, the APs are equipped with 35° directional antennas with 32 directions and an omni-directional antenna, and the clients are equipped with 4 sectors of 35 degree fan beam patch antennas and an additional omni-directional antenna. The reason for this setup will be explained in Chapter 5. For six transmissions, the upper bound network capacity is $40 * 6 = 240Mbps$, which is shown as the dotted line in the figures.

2.3.1 Q1: What are the benefits of power control?

By enabling power control (OptPC), the network capacity for *omn-tx,omn-rx* case can be improved by 31% in Campus 2 scenario and 70% in Lab 2 scenario. The benefits of power control for directional antennas systems decrease with stronger directionality. For example, for *dir-tx,dir-rx* case, power control only improves capacity by 4% and 7% in the two scenarios. This is because the network capacity of *dir-tx,dir-rx* case is already very close to the upper bound (dotted line), and leaves little space for improvement.

2.3.2 Q2: What are the benefits of directional APs or directional clients?

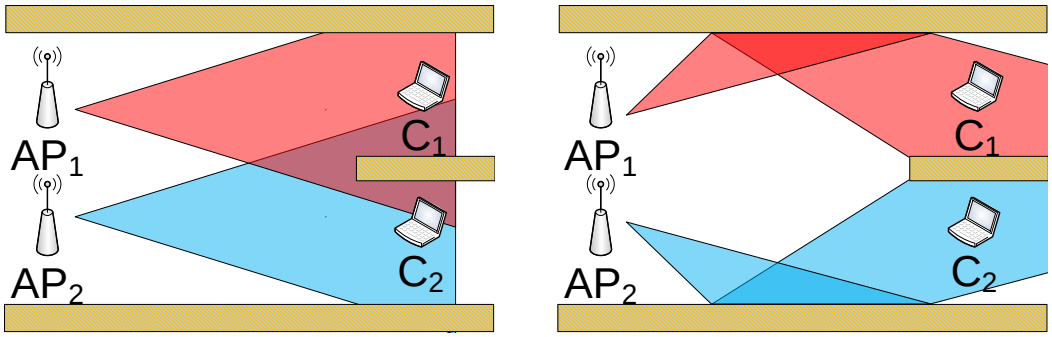
By deploying directional antennas on the APs, the network capacity can be improved over omni-directional antenna networks even with power control, by 30% and 51% in the two scenarios. When power control is also enabled for wireless networks with directional APs, the improvement is even higher, i.e., 41% and 71% respectively.

In fact, as described later in this section, the fundamental limitation of power control (compared with directional transmission) is that by decreasing the power level to decrease the interference level at unintended receivers, the signal level at the intended receiver is reduced by the same amount. On the other hand, by using directional transmission, the interference level at the unintended receivers can be significantly reduced without reducing the signal level at the intended receiver. Thus directional transmission is fundamentally more powerful than power control.

Since the directional antennas on the clients have weaker directionality than that on the APs (4 sectors of 35° fan beams patch antennas on clients vs. 32 directions of 35° antennas on APs), the network capacity of *omn-tx,dir-rx* case is a bit worse than that of *dir-tx,omn-rx* case.

2.3.3 Q3: What are the benefits of directional APs and clients?

By deploying directional antennas on both APs and clients, the capacity of the *dir-tx&dir-rx* configuration is close to the upper bound, i.e., all APs can transmit at 54 Mbps with reasonable frame loss rate. The reason is that the isolation provided by the directional antennas on both ends is higher than the SINR threshold to decode the 54Mbps frames. Also the configuration of *dir-tx&dir-rx* improves over the configuration of *dir-tx&omn-rx* by 38% in Campus 2 scenario, and 61% in Lab 2 scenario.



(a) Using MaxSNR directions (b) Exploiting reflected paths

Figure 2.4: Example of exploiting multiple paths using directional antennas (MaxSNR vs. MaxCAP)

2.4 Naive Solutions Do Not Work

The measurement based studies presented in the previous section show that directional antennas and power control can be very effective even in rich scattered indoor environments. In this section, we discuss why the naive solutions do not work well. We specifically look at three questions: 1) why uncoordinated solutions are unable to maximize spatial reuse, 2) why naive coordinated solutions are impractical, and 3) why CSMA based solution cannot exploit the spatial reuse opportunities.

2.4.1 Q1: Does MaxSNR work well?

MaxSNR antenna orientation approach, i.e., that the sender and the receiver choose antenna orientations that maximize the signal strength at the receiver, ensures that its transmission can be received at maximum signal strength. However, this approach only works well in isolation. If multiple directional senders exist in an indoor space, and can potentially transmit simultaneously, then orienting the senders according to the MaxSNR direction will not necessarily lead to the maximum spatial reuse, or system-wide capacity.

Figure 2.4 illustrates why the MaxSNR approach may not always maximize spatial reuse. In this example, nodes AP_1 and AP_2 are two directional senders that wish to transmit data to omni-directional receivers C_1 and C_2 respectively (*dir-tx,omn-rx* configuration). Given that there are no obstructions between senders and receivers, the MaxSNR direction is the same as the LOS direction (Figure 2.4(a)). Unfortunately, the LOS/MaxSNR directions lead to high interference at the receivers. For example, the transmission of AP_2 to C_2 interferes with reception at C_1, since C_1 is within the transmission range of AP_2. In this configuration, the MAC protocol must ensure that the two senders never transmit at the same time. In contrast, if the two senders select the orientations shown in Figure 2.4(b), then both senders could transmit simultaneously. Interference will still exist at the receivers, but it will be weaker, leading to a higher SINR at C_1 and C_2, and potentially successful packet receptions. This example illustrates that the best antenna orientation is not just a function of the receiver's location, but also a function of the

location of interfering transmitters. Note that the rich scattered indoor environment is leveraged in this example, i.e., reflectors that can potentially provide multiple paths between the sender and the intended receiver, and the obstructions that can potentially block interference between the sender and the unintended receiver.

In Chapter 4, we use measurement based studies to show that for the antenna configuration of directional APs only (*dir-tx,omn-rx*), MaxSNR does not work well in practice: MaxCAP outperforms 156% over MaxSNR in Campus 1 scenario, and 34% in Lab 1 scenario. In Chapter 5, we use measurement based studies to show that for the antenna configuration of directional APs and clients (*dir-tx,dir-rx*), though MaxSNR approach can achieve reasonable performance in two scenarios (in Campus 2 and Lab 2), it does not work well in other typical topologies, i.e., the performance of MaxSNR is worse than MaxCAP from 27% to 46% in Topo 1-5.

The results suggest that carefully configured directional antennas (MaxCAP) with OptMAC can significantly improve over the performance of the simple local heuristics (MaxSNR), though it may prove difficult to find this optimal configuration.

2.4.2 Q2: Does MinPC work well?

Using minimum power levels necessary for the receivers to decode (MinPC) will ensure that interference levels from that sender towards all the unintended receivers are reduced. However, MinPC does not in fact maximize spatial reuse because at the same time, the signal strength at the intended receiver is also minimized, thus leaving very little space to accommodate external interference.

Figure 2.5 shows an example of power control on omni-directional APs and clients (*omn-tx,omn-rx*). In this example, when APs deploy MinPC, AP_1 will use very low power level to C_1 and AP_2 will use high power level to C_2. In this case, no spatial reuse is possible because the signal level on C_1 is just enough to receive frames from AP_1 but the interference level on C_1 from AP_2 is too strong. In Chapter 6, we will more formally compare the NoPC, MinPC, and OptPC approaches using the SINR model, and the conclusion is that for NoPC, the transmission with longer distance (or higher path loss) is more likely to be penalized because its received signal strength is relatively weak. Using MinPC, however, has the opposite problem. If the sender is close to its receiver, it is more likely to be penalized because the interference level is likely to be high.

Also in Chapter 6, we use measurement based studies to show that the performance of MinPC is worse than OptPC from 21% to 103% across five different scenarios in practice.

The results suggest that carefully tuned power levels on senders (OptPC) with OptMAC can significantly improve over the performance of the simple local heuristics (NoPC and MinPC), though it may be difficult to find the optimal configuration.

2.4.3 Q3: Does CSMA work well?

CSMA for Directional Transmission

The most important assumption made in the CSMA protocol design is that the carrier sensed at the sender is a good indicator of the interference level at the receiver side. Two exceptions to this

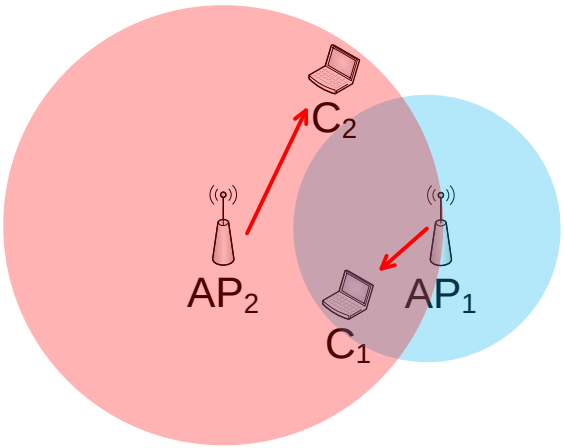
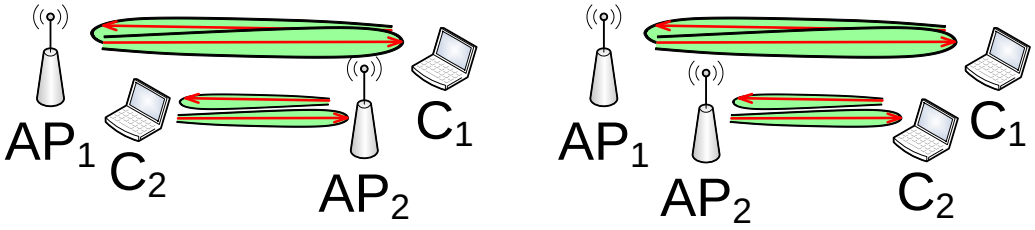


Figure 2.5: An example of power control: why MinPC or CSMA does not work well



(a) Directional exposed terminal

(b) Directional hidden terminal

Figure 2.6: CSMA in directional antenna networks

assumption are the well-known hidden terminal and exposed terminal problems. While CSMA works reasonably well in omni-directional antenna networks ([23] shows that the assumption is violated only in a very small number of scenarios), the biggest problem of using CSMA in directional antenna networks is that the number of violations of the assumption is significantly increased [108].

Figure 2.6 shows an example of directional exposed terminal problem and a directional hidden terminal problem for directional APs and clients (*dir-tx,dir-rx*). In Figure 2.6(a), the two APs can carrier sense each other because they are facing towards each other, but the interference level on both clients will be low because both clients are facing away from the interference sources. In Figure 2.6(b), the two APs cannot carrier sense each other because they are facing away from each other, but the interference level on both clients will be high because both clients are facing towards the interference sources.

Also in Chapter 4, we use measurement based studies to show that for directional AP systems (*dir-tx,omn-rx*), the performance of the default CSMA MAC protocol is very poor in practice: worse than that of the OptMAC by 108% in Campus 1 and 82% in Lab 1 scenarios, even with the MaxCAP antenna orientation algorithm.

CSMA for Power Control

The biggest problem of applying the default CSMA approach is link asymmetry, i.e., one sender can carrier sense another sender but not the other way round. We again use the example in Figure 2.5 to illustrate why the default CSMA does not work well in power control systems. In this particular example, no matter what power level is used on AP_2, AP_1 will always carrier sense and defer to AP_2, because AP_2 is far away from its intended client C_2, thus even if AP_2 uses minimum power to reach C_2, it will cause strong carrier sensing level at AP_1. There are two possible cases for AP_1, AP_1 can either use a low power level that causes link asymmetry (as shown in the figure), or it can use a high power level that removes the link asymmetry but at the same time, this is no spatial reuse possible.

Also, in Chapter 6, we use measurement based studies to show that for power control, the performance of the solutions that are based on default CSMA MAC protocol is worse than that of the OptMAC from 25% to 350%. CSMA with CCA tuning can alleviate several performance issues presented, but it is still not a desirable choice because the difference in performance from the OptMAC can still be up to 35%. Furthermore, starvation is possible in CCA tuning mechanism, and fine-grained CCA tuning mechanism cannot be implemented on existing hardware.

In summary, CSMA based solutions interact poorly with both directional transmission and power control and thus they do not work very well for either technique.

2.4.4 Q4: How about MaxCAP and OptPC?

The greatest drawback of both MaxCAP and OptPC approaches is efficiency: both require exhaustive measurements.

For MaxCAP, it requires exploration of all possible orientations of every sender and every receiver. The size of this search space grows exponentially when potential interferers and receivers are also directional, i.e., $O(K_{AP}^N \times K_C^N)$. As directional antenna technology improves, beam widths are likely to become smaller [6], thus increasing K_{AP} and K_C. This will render a brute force approach even more impractical.

Similarly, OptPC requires an exhaustive measurements of all possible power levels of every sender. Thus the size of the search space is exponential to the number of senders, i.e., $O(P_{AP}^N)$, which is simply too slow and impractical.

2.5 Motivating Three Systems

In previous sections, we have shown that directional transmission and power control can be effective in real world indoor environments, but naive solutions will not suffice to exploit these benefits. In this section, we motivate our three proposed systems. Based on the measurement studies, we characterize the design space for systems that optimize spatial reuse as a two-dimensional space: antenna configuration (omni-directional nodes, directional APs, and directional APs and clients) and the MAC protocol (centralized vs. distributed). This two-dimensional space, in fact, is a subcategory of the most top right point (coordinated and non-CSMA MAC layer) in the two-dimensional space in Figure 1.2.

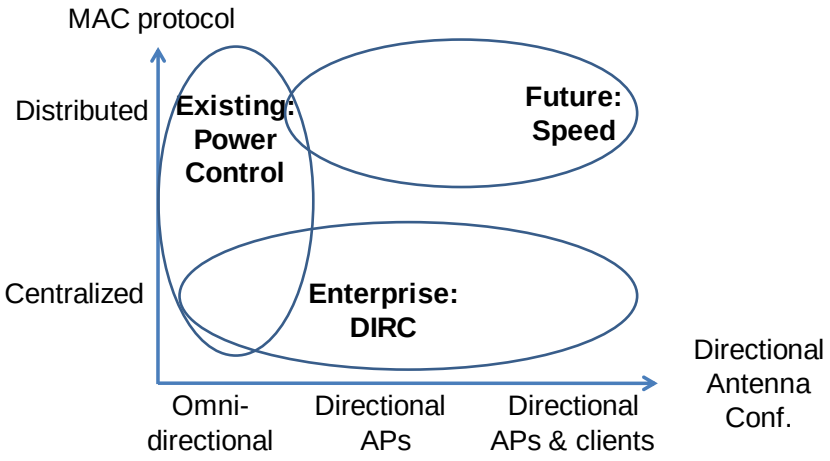


Figure 2.7: Design space for indoor spatial reuse

Figure 2.7 shows the design space. Note that we omit the directional clients on the dimension of antenna configuration because it is usually easier to deploy directional antennas on APs, especially directional antennas with stronger capabilities. Given the observations we made in this section, we choose three systems (three points) in the design space.

- For future wireless networks where both the APs and the clients can perform beamforming, we propose the Speed system which is a distributed system that exploits these capabilities. In fact, we also propose a practical antenna setup on wireless clients in Speed.
- For enterprise networks where directional antennas can be deployed on the APs and the clients still use omni-directional antennas, we propose the DIRC system, which is a centralized protocol that relies on the fact that the APs are connected through a separate wired connection.
- For existing omni-directional antenna networks, we propose the Opera system, which is a distributed power control protocol to improve spatial reuse for these networks.

We also believe the idea and the design of the three systems can be used to cover the other points in the space, as illustrated in Figure 2.7. The Opera system can be applied to both centralized and distributed MAC protocol. The Speed system can be applied to directional APs only with distributed control as well, though with a much lower expectation on performance. The DIRC system can be applied to all directional antenna networks with centralized control. This way, the whole design space is covered.

2.6 Related Work

There is also prior work on measuring how well directional antennas or power control works in indoor environments. The majority of the work has been focused on characterizing signal strength at the receivers. For example, early work [89, 129] has shown that in indoor environments, the signal strength from an AP to different locations is far from isotropic and does not have a strong relationship with the physical distance.

Recent work [20] has studied the performance of directional antennas in indoor environments, including directionality, spatial reuse, and localization. The authors observe that even though directionality is weaker (i.e., the resulting radiation pattern is less directional) in indoor environments, directional antennas can still offer enough isolation to improve performance over omni-directional antennas. They also observe that the LOS direction does not always have the strongest signal. The paper, however, only uses an indirect metric of reduced signal strength to evaluate spatial reuse, while in this dissertation, we consider network capacity. Also the work does not consider how to make use of directional antennas in indoor environments, even though their measurement results suggest that it is more challenging than in LOS environments.

In [24], the authors measure the performance of power control in wireless networks, and identify three cases: 1) overlapping, where aggregate throughput cannot be increased by power control, 2) hidden-terminal, where power control can help to ensure fairness, 3) potentially disjoint, where power control can allow concurrent transmissions. Their results are consistent with our observations. However, their work does not consider the implications of different MAC protocols.

Chapter 3

SINR Model Made Practical

In the previous chapter, we show that directional transmission and power control can be effective in indoor environments, when the nodes make coordinated decisions to optimize network-wide spatial reuse and use non-CSMA based MAC protocols. We also show that the naive coordinated solutions of exhaustive measurements are simply too slow and impractical. In this chapter, we present the key model to reduce the number of measurements in these coordinated solutions: the signal to interference and noise ratio (SINR) model and the conflict graph. We also present two assumptions to simplify our presentation and several important modifications to the original models that are critical in practice, e.g., incorporating multiple data rates, calibration, etc. The modified models are the core in all three systems (DIRC, Speed, and Opera) and the separation metric presented in this dissertation. At the end of the chapter, we also discuss related work on the interference models.

3.1 Interference Models

In this section, we introduce the interference models we use in this dissertation: the SINR model and the conflict graph.

As mentioned in the previous chapter, nodes in our systems make coordinated decisions on antenna orientations and power levels, and the SINR model is the key to significantly reduce the number of measurements necessary to make these coordinated decisions. The SINR model [42, 49, 64, 81] states that whether a frame can be successfully decoded at the receiver depends on the signal to interference plus noise ratio: if it is larger than a threshold, then the frame can be decoded, otherwise it cannot be decoded. Formally, the SINR model is:

$$\frac{Signal}{Interference + Noise} \geq SINR_{thresh} \text{ (in watts)}$$

where $SINR_{thresh}$ is the threshold, which depends on the transmission rate.

Conflict graph [10, 22, 53, 121] is another popular graph based tool to concisely encode the interference information. Vertices in conflict graphs indicate transmissions. For omni-directional systems, each vertex in conflict graphs is a transmission of (AP_i, C_j). The definition for indoor directional conflict graph is slightly different: each vertex is a virtual transmission, or the tuple

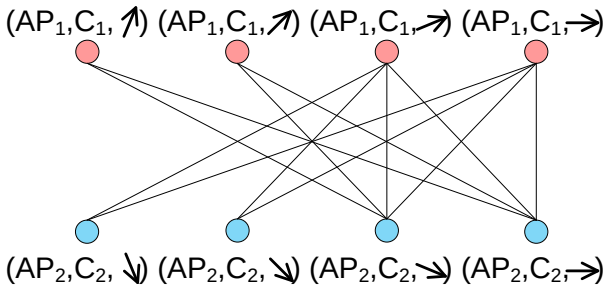


Figure 3.1: Example of an indoor directional conflict graph

of $(AP_i, C_j, K_{AP_i}, K_{C_j})$. Note that in the directional conflict graph defined in [116], each vertex is still the tuple of (AP_i, C_j) because the system is used for outdoor applications where there is only one direction for AP_i to reach C_j. While for indoor directional antenna systems, there usually exists multiple paths between the sender and the receiver, e.g., Figure 2.4. Edges in conflict graphs encode interference information: if there is an edge between two vertices, then the two (virtual) transmissions do interfere with each other and should not be allowed concurrent transmissions. On the other hand, the lack of an edge indicates that two (virtual) transmissions do not interfere and should occur simultaneously.

Figure 3.1 shows the partial indoor directional conflict graph for the network of Figure 2.4 with two transmissions. Since the full conflict graph for this network is very large, the figure only shows a small part of it. Note that the vertices do not include K_{C_j} because the clients receive omni-directionally. In this case, each vertex is a virtual transmission of the form of (AP_i, C_j, K_{AP_i}). The lack of an edge between vertices $(AP_1, C_1, \nearrow)$ and $(AP_2, C_2, \searrow)$, indicates that the two transmissions can occur simultaneously using direction $\nearrow$ and $\searrow$ on AP_1 and AP_2 respectively. On the other hand, the edge between vertices $(AP_1, C_1, \rightarrow)$ and $(AP_2, C_2, \rightarrow)$ indicates that the two transmissions cannot happen concurrently using the MaxSNR direction $\rightarrow$ on both senders. Note that we assume there are always implicit edges between vertices with the same AP or the same client, even though they are not shown in the figure. For example, there is an implicit edge between the vertices of $(AP_1, C_1, \nearrow)$ and $(AP_1, C_1, \rightarrow)$.

In the rest of this chapter, we will present several important modifications to make the models more practical or easier for presentation.

3.1.1 Noise Level and Pair-wise Assumption

In this section, we present two simplifying assumptions we made in the SINR model. It is worth noting that these two assumptions are made primarily to simplify the presentation and are not fundamental in our systems.

First, the thermal noise level in typical environments is around -95dBm (or 0.3pW), and since our work targets dense and interference dominated wireless networks, i.e., interference levels from other senders are much higher than -95dBm. We make the first simplifying assumption that thermal noise level is negligible. Also, in the rest of the thesis, we assume that signal strength, interference, and thresholds are all in decibel (dB), thus the SINR model now becomes:

$$Signal - Interference \geq SINR_{thresh} \text{ (in dB)}$$

The second simplifying assumption is the pair-wise assumption: that the interference level is dominated by the strongest source of interference. The pair-wise assumption is made implicitly in the conflict graph; making this assumption allows us to use the concise conflict graph. Prior work [36] evaluated the pairwise assumption and found it reasonable. It is worth noting that this assumption is made in this dissertation primarily because we use the conflict graphs to simplify our presentation and is not fundamental in our system design.

3.1.2 Multiple Data Rates

The SINR model and the conflict graph are only defined over a single data rate, but the notion of interference or successful reception depends on the data rate used. For example, the receiver may not be able to decode 54Mbps frames but can receive 36Mbps frames. Since supporting for multiple data rates is one of the most important features of today's wireless networks, here we define the SINR model and the conflict graph that incorporate multiple data rates.

For the SINR model, it can be easily extended such that there is one SINR threshold for each data rate: For some data rate $dr \in DR$,

$$Signal - Interference \geq SINR_{thresh,dr}$$

The key observation here is that given the signal and interference levels (or the SINR level), we can determine the highest data rate that can be supported on that link:

$$drm = \max_{dr}(Signal - Interference \geq SINR_{thresh,dr})$$

For the conflict graph, instead of each edge indicating whether two transmissions interfere or not, we annotate each edge by the data rate they can support on both transmissions. For example, for two vertices $(AP_{i_1}, C_{j_1}, K_{AP_{i_1}}, K_{C_{j_1}})$ and $(AP_{i_2}, C_{j_2}, K_{AP_{i_2}}, K_{C_{j_2}})$, first the SINR levels on both transmissions are computed as:

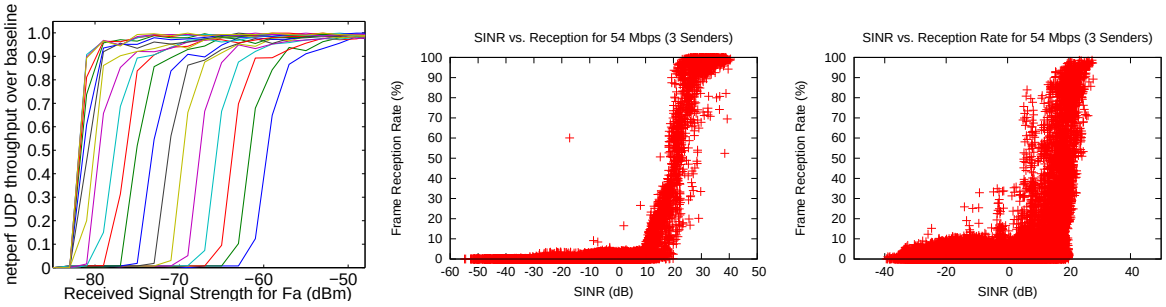
$$SINR_1 = S(AP_{i_1}, C_{j_1}, K_{AP_{i_1}}, K_{C_{j_1}}) - S(AP_{i_2}, C_{j_1}, K_{AP_{i_2}}, K_{C_{j_1}})$$
$$SINR_2 = S(AP_{i_2}, C_{j_2}, K_{AP_{i_2}}, K_{C_{j_2}}) - S(AP_{i_1}, C_{j_2}, K_{AP_{i_1}}, K_{C_{j_2}})$$

In this dissertation, we also call the two pair-wise SINR values PWS_1 and PWS_2. Then the edge between the two transmissions are annotated by the highest data rates that can be supported on the two transmissions, respectively, i.e., drm_1 and drm_2. Note that here both drm_1 and drm_2 have taken into consideration of protocol overhead for the data rates.

Given the annotated conflict graph, we can compute the capacity given a set of distinct virtual transmissions (that no two virtual transmissions has the same AP or client) T:

$$\sum_{(AP_{i_1}, C_{j_1}, K_{AP_{i_1}}, K_{C_{j_1}}) \in T} (\min_{(AP_{i_2}, C_{j_2}, K_{AP_{i_2}}, K_{C_{j_2}}) \in T} drm_1)$$

This formula says for each virtual transmission in T, we calculate the link throughput as when this virtual transmission is interfered by the strongest interfering transmission in T (or equivalently, the minimum drm_1). Then the capacity is the sum of all these link throughputs. This process takes $O(N^2)$ time.



(a) Emulator (b) Campus 1 (c) Lab 1

Figure 3.2: Accuracy of the SINR Model

3.2 Accuracy of the SINR model

In this section, we first examine how well the SINR model works in practice. Then based on the observations, we refine the SINR model accordingly.

Figure 3.2 shows how well the SINR model works in the wireless emulator [59], Lab 1, and Campus 1 scenarios. In the emulator experiment, we use two pairs (i.e. four) of laptops. We use *netperf* [58] to create a single non-rate controlled UDP flow on each of the two node pairs, namely F_a and F_b, with transmit power levels P_a and P_b, respectively. The power levels are set manually. Also, we use the emulator to prevent any interference at both sources - thus eliminating the effects of carrier sense and collisions with link-layer ACKs. Figure 3.2(a) shows throughput achieved by F_a. Ideally, we expect the curves to be threshold functions. We keep P_b constant and vary P_a for each curve and vary P_b between curves. The well-spaced and similar-shaped curves, with an exception when thermal noise dominates interference, show that SINR model is reasonably accurate. The other two experiments are carried out in Campus 1 and Lab 1 scenarios, where we use three APs as concurrent senders. Notice that the same set of hardware (wireless cards) is used in the Campus 1 and Lab 1 scenario, but it is different from the hardware used in the emulator experiment. The details of the experimental setup is described in Chapter 4. Figure 3.2(b)&(c) show the scatterplots of SINR and frame reception rates. The shapes of the curves suggest that the SINR model is reasonably accurate in practice. Other work [70] has also made similar observations. Next we will look at several issues with the original SINR model.

3.2.1 Capture Effect

In fact, the validity of the SINR model is dependent on hardware implementation, or an effect called the capture effect [70, 80, 103, 123]. The capture effect is the ability for the receiver hardware to pick up a second frame with sufficiently strong signal to interference ratio even when the receiver is actively receiving the first frame. If the capture effect presents on the hardware, the SINR model is more accurate; otherwise, the simple range based model (described in Chapter 3.3) is a better choice. All the wireless cards with Atheros [2] chipsets we used in our experiments support the capture effect, thus the SINR model is valid for these cards. Even though not all cards support the capture effect, there are other techniques such as raising the

receiver threshold, to make the SINR model work for cards that do not inherently support the capture effect [123].

3.2.2 Intermediate Loss Rates

Another observation made both in previous [9] and in our work (Figure 3.2) is that there are many data points with an intermediate loss rates, even in the wireless emulator without the presence of multipath effects. In order to incorporate the intermediate loss rates, we modify the SINR model to have two thresholds $SINR_{high}$ and $SINR_{low}$ such that if the SINR is lower than $SINR_{low}$, few frames can be decoded, if the SINR is higher than $SINR_{high}$, almost all frames can be decoded, and if the signal level is between the $SINR_{high}$ and $SINR_{low}$, the reception rate is intermediate and is linear to the SINR level. Formally, the packet reception rate (PRR)

$$PRR = \begin{cases} \alpha & \text{if } SINR > SINR_{high} \\ \beta & \text{if } SINR < SINR_{low} \\ \frac{SINR - SINR_{low}}{SINR_{high} - SINR_{low}} & \text{otherwise} \end{cases}$$

Here, α is a number close to 1 and β is a number close to 0. In our work, we set $\alpha = 0.9$ and $\beta = 0.1$. This is because according to Figure 3.2, it is rarely the case that $\alpha = 1$ and $\beta = 0$ in practice.

Calibration

The SINR model relies on accurate SINR thresholds on receivers to work well. Unfortunately, previous work [43, 59] and our work has observed that different wireless cards may have different SINR thresholds. For example, difference of up to 10dB have been observed in [59], and up to 4dB has been observed in our work (Chapter 6). Even if the devices with the same manufacturer can be calibrated before shipping out, the devices can become uncalibrated over time [43].

In our SINR model, both $SINR_{high}$, $SINR_{low}$ need to be calibrated. Generally speaking, the thresholds are dependent on each individual card and the multipath channel condition between the sender and the receiver. The reason that the multipath channel condition may matter is because the SINR value is a single value computed over all the sub-carriers of the OFDM spectrum. In an OFDM frame reception, multipath effects will cause low signal strength in several sub-carriers in the OFDM spectrum. Thus different multipath channel conditions may result in different patterns of signal strength in the OFDM sub-carriers and different receiving patterns but similar SINR values. There is the tradeoff between calibration accuracy and calibration overhead. Ideally, the calibration results are the most accurate if it is done on each pair of vertices in the conflict graph. However, this is simply too slow and impractical. On the other hand, the minimum calibration requirement is per-environment calibration, as Figure 3.2 illustrates that the thresholds may differ even with the same wireless cards in different environments, i.e., the range of intermediate loss rates is larger in the Lab 1 scenario than that in the Campus 1 scenario. In our measurements, we find that for DIRC and Speed, the per-environment calibration results is good enough and for Opera, a per-card calibration is necessary.

While manual calibration using a signal analyzer and generator is possible [43], this is not a practical solution in most deployments. In Chapter 4 and Chapter 6, we present the details of the automatic calibration on a per-environment and a per-card basis, respectively.

3.3 Related Work

Modeling Interference

Most prior work on power control implicitly uses a simple range-based model [11, 95]. In this model, every source is associated with a transmission and an interference range. Nodes within the transmission range of a source can decode frames from the source, and nodes within the interference range will be prevented from transmitting due to carrier sense [22]. The ranges depend on the power level each source uses. The range-based model predicts that using the minimum possible power level minimizes interference. On the other hand, the SINR model predicts that uniformly increasing power levels increases system throughput by reducing the effects of thermal noise, though the gains are marginal in interference-dominated networks, e.g. dense 802.11 networks. Interestingly, this is exactly the opposite conclusion of the range-based model.

The right choice of model, in fact, depends on whether the interface hardware exhibits the capture effect. If the interface can capture the stronger signal, then the SINR model is more accurate – otherwise the range-based model is a better predictor. While in this chapter, we show in experimental results that existing hardware exhibit strong capture effect, indicating that the range-based model is not representative for existing hardware. Similar observations were also made in [70, 80, 103], and the SINR model is also widely accepted in many research efforts [22, 42, 49, 64, 81]. A more detailed explanation and comparison can be found in our earlier work [74].

Directional Conflict Graph

In [116], the authors discuss the definition and generation of directional conflict graphs, with additional consideration of side lobes, but the primary usage is in outdoor ad-hoc networks with little or no multipath. Thus the definition of directional conflict graph is different from the indoor directional conflict graph defined in this chapter. Basically, each vertex in an outdoor directional conflict graph represents an actual transmission of (AP_i, C_j); while each vertex in an indoor directional conflict graph represents a virtual transmission of $(AP_i, C_j, K_{AP_i}, K_{C_j}$. The difference is due to rich scattering in indoor environments, i.e., multiple paths exist between senders and receivers.

Generating the Conflict Graph

The conflict graph model have also been widely used in prior wireless network systems [10, 22, 53, 121]. In prior efforts, however, conflict graphs are constructed through measurements of link throughputs, both synchronously and asynchronously. For example, [10] proposes to use centralized and synchronized throughput micro-benckmarks to quickly construct the complete

conflict graph. In [121], the authors propose to construct the conflict graph by detecting packet collisions (an asynchronous approach). In our work, we construct the conflict graphs using the SINR model. The biggest advantage of these throughput based solutions to construct the conflict graph is high accuracy: it does not suffer from SINR calibration errors, and the synchronous approach does not suffer from undecodable frames. However, the biggest problem with these approaches is that they require a large number of measurements to construct the conflict graph:

- Throughput based solutions require pair-wise throughput measurements, thus the number of measurements is square of the number of virtual transmissions; while SINR based solution only requires the number of virtual transmissions worth of measurements.
- For directional antenna networks, each actual transmission has multiple virtual transmissions, i.e., one for each antenna orientation, for both throughput and SINR based solutions.
- For power control, each actual transmission in throughput based solutions has multiple virtual transmissions, i.e., one for each power level; while for SINR based solution, each actual transmission has only one virtual transmission.
- For any system that supports multiple data rates, each actual transmission in throughput based solutions has multiple virtual transmissions, i.e., one for each data rate; while for SINR based solution, each actual transmission has only one virtual transmission.

Thus the number of measurements for throughput based conflict graph generation is $O((N \times K_{AP} \times P_{AP} \times DR)^2)$, while that for SINR model based generation is only $N \times K_{AP} + M$ (See Section 2.1 for notations). This problem prevents the throughput based solutions from being effective in directional antenna and/or power controlled networks. Note that for omni-directional antenna networks without power control, these solutions can actually be quite desirable because the number of measurements $O((N \times DR)^2)$ is still manageable [10].

Aside from the number of measurements, as shown in Chapter 4, the measurements of the SINR values are much more stable than that of the throughputs. Thus much less time is needed to accurately measure the SINR values than throughputs (5ms vs. 100ms).

In our systems, we find that the conflict graph generated using the SINR model is accurate enough for our purposes, with several techniques such as calibration and feedback mechanism. Note that for other applications where measurement overhead is manageable, e.g., for omni-directional antennas [10, 109], throughput based generation is a better option.

Interference Cancellation

Recent work on interference cancellation [46, 47, 50] demonstrates that with smart signal processing techniques, the SINR requirement (that the signal strength of the desirable frames must be sufficiently stronger than that of the undesirable frames) may not be necessary. For example, [46] shows that with ZigZag decoding, a receiver may be able to decode two frames even when the signal strength from both frames are similar. In fact, ZigZag decoding works the best when the signal strength from two frames are similar. Note that ZigZag decoding still requires two collisions. When the signal strength from the two frames are significantly different, [50] proposes a mechanism to receive both frames by first cancelling out the frame with stronger signal. These interference cancellation mechanisms may have strong implications on how hidden terminal problems should be handled in future wireless systems, and in our proposed systems.

For example, it may be desirable to have directional senders either to achieve similar SINR levels at various receivers, or to allow the interference level to be sufficiently strong and keep signal level low. However, current research in interference cancellation is still in an early stage: 1) most experiments focus on two senders, and how these techniques can handle multiple senders is unclear, and 2) current experiments are carried out using simple modulation schemes and MAC protocols in GNURadio [4] and how well these techniques work for more complicated 802.11 systems is unclear. Also, even when these solutions become popular, the framework of our solutions remains the same: using the SINR model to characterize the environment the wireless nodes operate in and then plug in different optimization algorithms on top of the collected measurements.

Chapter 4

DIRC: Directional AP System for Enterprise Networks

In Chapter 1, we show that the demand for wireless bandwidth in indoor environments continues to increase with the rapid integration of Wi-Fi radios in every day consumer electronic devices. This increased density of Wi-Fi transmitters has exacerbated contention for the wireless medium and reduced overall throughput. One effective approach to improving throughput is to reduce interference among radios, thus allowing them to transmit concurrently.

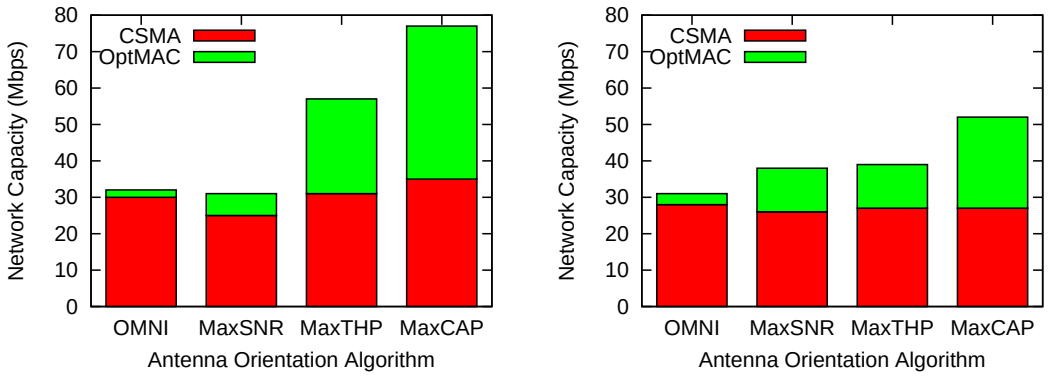
In Chapter 2, we show that directional antennas have the potential to provide the necessary interference reduction by spatially confining transmissions. For example, commercial, off-the-shelf directional antennas (Figure 2.1) can provide spatial isolation of 10-15dB by confining the signal within a sector, i.e., the signal outside the sector is at least 10-15dB weaker than the signal within the sector. This high degree of spatial isolation can support concurrent transmissions between pairs of radios—provided that the antennas can be oriented correctly.

Traditionally, directional antennas have mostly been used in outdoor environments where there is a direct line-of-sight (LOS) between the two endpoints [90, 97]. In these applications, orienting directional antennas is relatively simple because the LOS antenna orientation usually maximizes the signal strength at the receiver and is the only configuration that works well.

In an indoor environment, the presence of rich scattering and multipath effects results in non-LOS antenna orientations that provide comparable signal strength at the receiver. In fact, conventional wisdom has been that directional antennas will not be effective indoors because of this issue. Although the existence of these alternate good orientations complicates the configuration of the system, it also creates an opportunity for reducing interference between transmissions.

However, to achieve the gains offered by these alternate configurations requires significant, explicit coordination between the transmitters in the network. For example, coordination is necessary for transmitters to measure the interference created in different antenna configurations. In addition, since the presence of directional antennas exacerbates the presence of hidden and exposed terminals in the network [108], coordination is also needed to address MAC related problems.

In this chapter, we present DIRC [75], a wireless network design that improves spatial reuse in indoor environments using directional APs. The DIRC design focuses on exploring a typical enterprise environment in which: 1) the access points (APs) are centrally controlled and man-



(a) Campus 1

(b) Lab 1

Figure 4.1: Median network capacity under different antenna orientation algorithms and MAC protocols

aged, 2) the APs are equipped with phased array (software steerable, directional) antennas, and 3) the clients have standard 802.11 omni-directional hardware. The core of DIRC is an algorithm that identifies close-to-optimal orientations for the directional antennas (the algorithm is based on the SINR model and the conflict graph presented in Chapter 3), maximizing system-wide capacity while ensuring that configuration overhead is low and scales linearly with the number of APs. DIRC also incorporates a new timeslot-based MAC protocol designed for indoor directional antennas. This design is much more efficient than CSMA/CA-based MAC protocols, which assume that the interference at the sender is similar to the interference at the receiver. This assumption breaks down dramatically with directional antennas (as presented in Chapter 2).

4.1 Opportunities and Challenges

In Chapter 2, we have shown that directional APs can be effective in two indoor scenarios (Campus 2 and Lab 2), but neither the MaxSNR approach nor CSMA can fully exploit these benefits (through examples in Figure 2.4 and Figure 2.6). In this section, we present additional measurement study in two other scenarios (Campus 1 and Lab 1) to examine the performance of all antenna orientation algorithms and MAC protocols across these two scenarios.

Figure 4.1 illustrates the dramatic performance impact that correctly orienting directional antennas can have. These experimental results are for two scenarios (Campus 1 and Lab 1), each using three directional transmitters. The testbed and experimental setup are described in Chapter 4.4 (this figure only shows median capacity, and mean capacity is similar). The antenna orientation algorithms of OMNI, MaxSNR, MaxCAP are the same as those are presented in Chapter 2, and we also introduce an alternative simple local heuristic of MaxTHP:

- OMNI: The performance observed when the APs transmit omni-directionally
- MaxSNR: The simple heuristic where each sender chooses the direction that has the maximum SNR, as observed at the receiver.

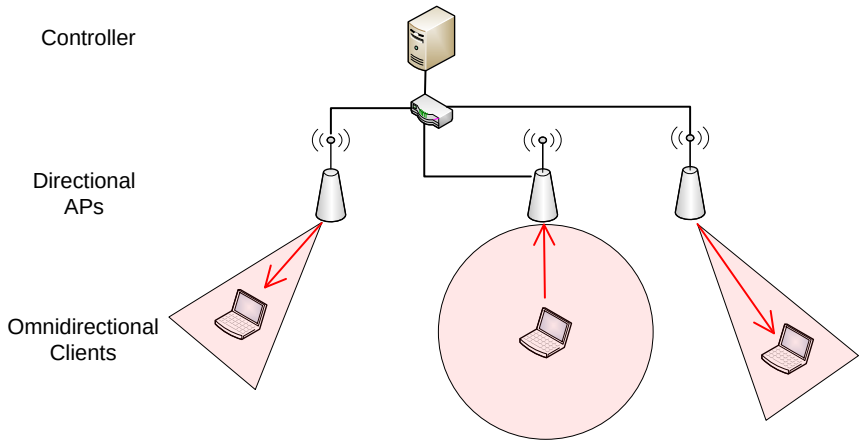


Figure 4.2: DIRC system overview

- MaxTHP: The heurstic where each sender chooses the direction that maximizes its individual link throughput. Note that this heuristic relies on an iterative process by which each sender picks the direction that maximizes its own throughput given the directions selected by the other directional senders. The final configuration of MaxTHP is either one where no sender can further improve its throughput or the one reached after 30 algorithm iterations.
- MaxCAP: The performance obtained by orienting the antennas in directions that maximize system-wide capacity, and the orientations are determined through exhaustive search. Note that the performance of MaxCAP to illustrate the maximum capacity that can be achieved in each scenario.

For each of these antenna orientation schemes, we operate the system using the 802.11 CSMA MAC protocol (red/lower bars) and using an optimal MAC protocol that is capable of coordinating transmissions across space while avoiding hidden and exposed terminals (OptMAC - red plus green/lower plus upper bars).

Figure 4.1 shows that MaxCAP with OptMAC can provide about twice the performance of using omni-directional antennas (also with OptMAC). Note that the maximum possible gain with three transmitters will be a factor of three. Unfortunately, simple local heuristics (MaxSNR and MaxTHP) are unable to identify configurations that perform close to optimal, even with OptMAC. The result shows that there can be significant performance gains. The graph also highlights the importance of an effective MAC protocol; the CSMA protocol, for example, does a poor job using the capabilities of the directional antennas.

In fact, the problem of antenna orientation becomes even harder when one considers RF environment variability and user mobility. Any algorithm for optimal antenna orientation across directional senders will need to be able to collect, process and act on information very quickly. Current electronically steerable antennas can change their orientation in $100\mu s$. The challenge is to design an algorithm that provides the best tradeoff between optimality and stability/overhead.

DIRC addresses both of these aspects: a close-to-optimal antenna orientation algorithm and a MAC protocol design. Next we show the design of DIRC, and later we will show that the performance of DIRC is very close to that of the MaxCAP with OptMAC.

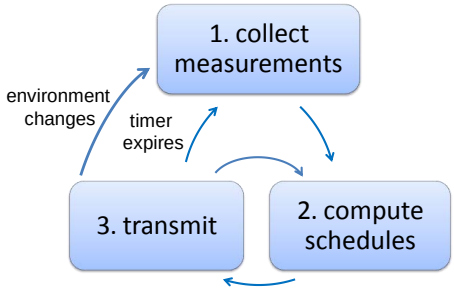


Figure 4.3: DIRC operations

4.2 DIRC Design

DIRC is designed and built for enterprise wireless networks where the APs have software steerable directional antennas with small reconfiguration times (e.g., phased array directional antennas), and all of the clients are omni-directional (e.g., laptops). The APs are connected to each other through a separate, independent channel such as wired Ethernet and are coordinated by a central controller. Figure 4.2 shows the system overview.

4.2.1 Notations

Since all the clients in DIRC are omni-directional, we use a simpler notation for signal strength than those used in Chapter 2: we use $S(AP_i, C_j, K_{AP_i})$ to denote the signal strength from AP_i to C_j with orientation K_{AP_i} on AP_i. The other notations are the same as presented in Chapter 2.

4.2.2 Overview

The core of DIRC's design is to use the SINR model and conflict graph presented in Chapter 3 to reduce the number of measurements needed to orient the antennas. In DIRC, only $N * K_{AP}$ measurements are necessary to implement the heuristics of the MaxCAP algorithm, which is the same number of measurements needed to implement the MaxSNR algorithm.

DIRC uses a timeslot based centralized MAC protocol. In order to prevent the interference from omni-directional transmissions, each timeslot is split into two phases, dirc-tx phase where only directional APs can transmit, and omni-tx phase where omni-directional clients can initiate transmissions. In dirc-tx phase, the centralized controller computes a schedule for the directional APs; while in omni-tx phase, the default CSMA MAC is used to coordinate the omni-directional transmitters. In fact, this split is not desirable in many scenarios due to its inflexibility. In DIRC design, we propose a simple re-split mechanism to provide more flexibility, a better and even more flexible solution, is presented and used in Speed, as described in Chapter 5.

Figure 4.3 shows the three stages of DIRC operation: 1) collecting measurement, 2) scheduling, and 3) transmitting. The measurements need to be updated on two conditions: first, the measurements need to be updated periodically, and second, when the environment changes dramatically.

In the measurement stage, the centralized controller instructs each AP to send a number of frames in each direction, and all receivers will record the signal strength from that AP with a particular direction. During this stage, the set of measurements collected is the complete table of $S(AP_i, C_j, K_{AP_i})$.

In the scheduling stage, the centralized controller will apply the SINR model on the collected measurements to determine which APs can transmit and the best antenna orientations for those APs. In DIRC, the centralized controller keeps a FIFO queue of backlogged transmissions, and then tries to include as many non-interfering transmissions to the next timeslot as possible. For each combination of antenna orientations, the controller uses the conflict graph to infer the network capacity as the sum of all link throughputs (as shown in Chapter 3). Finally, it picks the best combination of antenna orientations and uses that as the schedule for the next timeslot.

In the transmission stage, the backlogged APs will inform the controller that they have frames to be transmitted to particular clients. Then the controller computes the schedule for the next timeslot and sends the schedules to the APs. Then in dirc-tx phase, APs will set their directions and send frames to the clients. In omni-tx phase, the clients can send link-layer ACKs and data frames to the APs. Both APs and clients need to be modified.

In the following sections, we present the details of the three stages.

4.2.3 Collecting Measurements

In Chapter 3, we show how to construct the conflict graph using the SINR model with multi-rate support. To construct the conflict graph, the necessary information is the complete table of signal strength from all the APs to all the clients.

We collect the necessary signal strength information as follows. Each directional AP scans all its directions and transmits a burst of frames in each direction (the duration of the burst is set to 5 ms, which is determined empirically in Chapter 4.3.1). All of the clients record the received signal strength indicator readings (or RSSI readings) for the frames received, and they calculate $S(AP_i, C_j, K_{AP_i})$ as the mean RSSI at client C_j from AP AP_i using direction K_{AP_i}. When one directional AP is scanning, clients that are not associated with this AP also record the received RSSI as the interference level from that AP. Throughout this process, we assume that clients can be modified to collect the above measurements. To ensure that the measurement probes can be received at the client, during each measurement, all APs are prevented transmitting except the scanning AP, and each AP is instructed to scan in turn.

One important limitation of our measurement process is that we only consider decodable frames, but undecodable frames can also interfere; note that throughput micro-benchmarks described in Chapter 3.3 do not suffer from this problem. To address this issue we proceed as follows. First, all probe frames use the minimum data rate (1 Mbps) to maximize the probability of decoding. Second, the clients report the RSSI readings of all frames even if they are erroneous. Note that readings from erroneous frames have to be treated very carefully and only used as secondary choice, since many readings are set to -95dBm. Then, if some $S(AP_i, C_j, K_{AP_i})$ is still unavailable (i.e., no frame has been received), we use the following criteria to determine its value:

- If client C_j can receive frames from AP_i using other directions, we set $S(AP_i, C_j, K_{AP_i})$ to the minimum decodable signal level, i.e., -95 dBm. The actual signal strength, in fact, will

be lower than this level (otherwise they can be decoded). The reason we take a conservative approach in this case is that the AP and the client are not too far away from each other (that the client can hear the AP from other directions).

- If no frames from AP_i can be received by client C_j at all, we set $S(AP_i, C_j, K_{AP_i})$ to $NO\_INTERFERENCE$. The reason we take an aggressive approach is that it is very likely that the AP and the client are very far away and there will be no interference from the AP to the client.

Such a solution may still fail in scenarios where the interference level from a couple of directions could be right below the decodable threshold, and may "hide" hidden terminal interference. In the next section, we discuss a feedback mechanism, aimed at dealing with these cases. Finally, we observed that the prediction of the link throughput $S(AP_i, C_j, K_{AP_i})$ is usually very poor if only a few frames are available for the calculation. Consequently, our algorithm treats measurements produced by fewer than 10% of all frames as if no frame was received.

After the measurements are collected at the centralized scheduler, it can construct the conflict graph, as shown in Chapter 3.

4.2.4 Scheduling Using the Conflict Graph

Given the conflict graph, the second phase of the system operation is for the central controller to determine which transmissions can take place and, if they are directional, with what orientation. Such a decision is made for each timeslot and based on the actual traffic pattern across the network.

Full Scheduling Problem

The scheduling problem to optimize capacity for directional antennas is NP-hard (the scheduling problem can be reduced to maximum clique problem). The size of the problem is $\binom{N}{N}K_{AP}^N + \binom{N}{N-1}K_{AP}^{N-1} + ... + \binom{N}{1}K_{AP}$, where each $\binom{N}{i}$ is the number of ways the controller can pick i out of N transmissions, and K_{AP}^i is the number of all possible antenna configurations for these i directional APs. Notice here that even the simpler, special case of omni-directional antennas where $K_{AP} = 1$, is NP-hard [14, 22, 105].

Reducing the Search Space in Scheduling

We propose two heuristics to reduce the complexity. Algorithm 3 shows this process for one time slot. First, we order the transmissions according to the time their frames arrived at the APs (i.e., FIFO). The scheduler goes through each transmission, t, in order and adds it to the schedule for the next time slot, T_{next}, if it improves the network capacity, $maxcap$. This step reduces the complexity to $K_{AP}^N + K_{AP}^{N-1} + ... + K_{AP}$. Any unscheduled transmissions will remain in the FIFO queue, and those transmissions that have been scheduled and have more frames to send will be appended at the tail. This way, starvation is prevented. In the worst case, a transmission will be scheduled after $N - 1$ time slots. At the same time, the channel will be well-utilized, since transmissions with fewer conflicts will be scheduled more often. Note that this approach of

Algorithm 1: Calculate link throughput given the SINR value ($calc\_thp$)

Input: SINR value $SINR$
Data: set of data rates DR, SINR thresholds on this link for each data rate $SINR_{high,dr}, SINR_{low,dr}$
Output: link throughput lt

$lt \leftarrow 0$
/ test each data rate dr on the SINR level */*
foreach $dr \in DR$ **do**
    / SINR very strong: no loss */*
    if $SINR > SINR_{high,dr}$ **then**
        $thp \leftarrow dr$
    / SINR very weak: no reception */*
    else if $SINR < SINR_{low,dr}$ **then**
        $thp \leftarrow 0$
    / otherwise: intermediate loss rates */*
    else
        $thp \leftarrow ((SINR - SINR_{low,dr})/(SINR_{high,dr} - SINR_{low,dr})) * dr$
    end
    / data rate dr yields high throughput */*
    if $thp > lt$ **then**
        $lt \leftarrow thp$
    end
end

Algorithm 2: Compute network capacity for a schedule, i.e., set of transmissions ($estimate\_capacity$)

Input: a set of transmissions T
Output: the network capacity cap

$cap \leftarrow 0$
/ For each transmission */*
foreach $(AP_{i_1}, C_{j_1}, K_{AP_{i_1}}) \in T$ **do**
    $sig \leftarrow S(AP_{i_1}, C_{j_1}, K_{AP_{i_1}})$
    $max\_intf \leftarrow -95$
    / Find the strongest interfering transmission */*
    foreach $(AP_{i_2}, C_{j_2}, K_{AP_{i_2}}) \in T$ **do**
        $intf \leftarrow S(AP_{i_2}, C_{j_1}, K_{AP_{i_2}})$
        if $intf > max\_intf$ **then**
            $max\_intf \leftarrow intf$
        end
    end
    / Sum up the link throughputs */*
    $cap \leftarrow cap + calc_thp(sig - max\_intf)$
end

servicing requests will not lead to more unfairness than what will be expected in a normal 802.11 network; senders that experience more conflicts will have fewer opportunities to transmit.

Second, we apply a greedy algorithm to find the orientations. Instead of visiting all K_{AP}^N states, the scheduler emulates an iterative process where all APs start at antenna orientation 1 and then taking turns to maximize the network capacity by varying the orientation of its own antenna and by keeping the orientations of other APs fixed. This iterative process will converge because in each iteration, the capacity is improved. The number of rounds for the algorithm to converge depends on individual scenarios, and in the worst case, the complexity of the greedy algorithm can be $O(K_{AP}^N)$. In our algorithm, we set an upper limit for 30 rounds, and, in practice, we find the algorithm converges after approximately 16.25 rounds on average. Note that we choose these two heuristics because they are simple and they perform well in our testbed. In a very large system, however, other heuristics may provide better performance. For example, algorithms like the one proposed in [22] for capacity maximization in omni-directional antenna networks, or simulated annealing may be better alternatives to meet the system goals and avoid local maxima. Indeed, we tested these two alternatives in our scheduler, but they did not provide further improvement.

Feedback

In some cases, the SINR model may fail to predict the link throughput, for example, when a client cannot decode frames from another interferer. Our system uses a feedback mechanism to deal with this problem. The idea of a feedback mechanism is that if a certain scheduling assignment

Algorithm 3: Pick directions and assign timeslot

Data: FIFO queue of active transmissions $N_{all} = \{t\}$
Output: assignment of next time slot and direction $T_{next} = \{(t, k)\}$, where t is the transmission and k is the direction to be used; note that unscheduled transmissions will remain in the queue

$T_{next} \leftarrow \emptyset$
$maxcap \leftarrow 0$
$N_{cur} \leftarrow \emptyset$
/ test if t can be added to T_{next} */*
foreach $t \in N_{all}$ **do**
    $N_{cur} \leftarrow N_{cur} \cup \{t\}$
    $T_{last} \leftarrow \emptyset$
    $T_{cur} \leftarrow \{(t, k)$, where $t \in N_{cur}$ and $k \leftarrow 1\}$
    / each AP chooses orientation in turn, greedily increasing network capacity until convergence */*
    while $T_{cur} \neq T_{last}$ **do** */* test if T_{cur} changes since last iteration */*
        $T_{last} \leftarrow T_{cur}$
        foreach $t' \in N_{cur}$ **do**
            $localmaxcap \leftarrow 0$
            $T_{tmp} \leftarrow T_{cur}$
            / find orientation of t' that maximizes capacity, with other orientations fixed */*
            foreach $k \in K_{AP}$ **do**
                $T_{tmp}[t'] \leftarrow (t', k)$ */* set the orientation of t' to k */*
                $curcap \leftarrow estimate\_capacity(T_{tmp})$
                if $curcap > localmaxcap$ **then**
                    $localmaxcap \leftarrow curcap$
                    $T_{cur} \leftarrow T_{tmp}$
                end
            end
        end
    end
    / if including t can increase network capacity, then it should be scheduled for next timeslot */*
    if $maxcap < localmaxcap$ **then**
        $maxcap \leftarrow localmaxcap$
        $T_{next} \leftarrow T_{cur}$
    end
end
$N_{all} \leftarrow N_{all} - T_{cur}$ /* remove scheduled transmissions */

fails to deliver the expected performance, the controller will pick the next best scheduling assignment. Note that it is possible that the prior assignment was indeed the optimal assignment but the actual measured throughput was lower than the estimated throughput. In our measurements, we observe that although there is some difference between the estimated and actual throughput, the estimation mostly preserves the ordering (i.e., schedules with higher expected throughput provide better throughput). Consequently, our algorithm samples at most three scheduling assignments and picks the best. In the greedy heuristic in Algorithm 3, we use a different starting state in each iteration, and could thus converge to a different final state.

4.2.5 Dynamics

A practical solution needs to account for the dynamics caused by people moving, doors opening and closing, and changes in the location of wireless devices (e.g., laptops). We identify two types of dynamicity: short-term dynamics caused by multipath fading and other short-lived environmental changes, and long-term dynamics caused by more permanent changes.

Short-term dynamics cause the RSSI and link throughput measurements to fluctuate quickly. The APs should not respond to these changes, since the coherence time is too short for our algorithm to rescan and reconfigure. Using a moving average is a common mechanism to obtain more stable and statistically sound readings.

Long-term changes, on the other hand, should trigger the regeneration of the conflict graph based on new measurements. In some cases, changes can be readily observed (e.g., loss of throughput on a link or packet losses due to collisions) while in other cases they cannot (e.g., a previously inefficient transmission schedule is now attractive because a door was closed). As a result, pure failure-based re-scanning is not enough, and we need to periodically rebuild the conflict graph (Chapter 4.3). The challenge is to find a good tradeoff between a rapid response to long-term dynamics while ignoring short-term effects. Note that periodic rescanning does not help with the SINR model mis-predictions discussed above, since these are caused by limitations of the node measurement capabilities. That is, the difference between SINR failures and dynamicity in the environment, is that dynamicity usually incurs changes in reception rates and signal levels but SINR failures do not. Thus if the frame reception rate is lower than previous observed and signal level drops, it is considered dynamicity. While if the reception rate is consistently low, it is considered SINR failures.

4.3 DIRC Implementation and Operation

Now we present the details of DIRC's protocol of operation. Our DIRC implementation is based on the Madwifi driver and Atheros 5212 wireless cards, present on both clients and APs. The controller is a Linux machine that can access all APs through the wired Ethernet.

4.3.1 Measurement Collection and Scanning

DIRC builds the conflict graph based on RSSI measurements that are collected across the network. To maintain an accurate conflict graph, it is important that the RSSI measurements adapt

Link	Win. Size (ms)	5	10	100	1000
Least Bursty	RSSI (dB)	0.2	0.16	0.05	0.03
	Reception (%)	25	21	15	5
Most Bursty	RSSI (dB)	2.2	1.5	1.3	1
	Reception (%)	50	35	26	16

Table 4.1: Allan deviation of RSSI and frame reception rate for different intervals

to short-term as well as long-term dynamics of the environment while minimizing the overhead on the system performance. Thus, DIRC periodically instructs all APs to scan for updated RSSI measurements, or a scan can also be explicitly requested by the APs when the measured throughput on any link drops below a set threshold.

The protocol between the APs and the central controller to initiate a scan is as follows. Scans are requested by the AP through a *request-to-scan* message. The controller initiates the scan by sending a *clear-to-scan* message to the AP; scans in response to AP requests are dampened and take place at most once a second. The controller gives higher priority to scan requests than to transmissions, since efficient transmissions require an accurate conflict graph. Each scan probe includes the AP identifier and the direction used for the probe frame. When the client receives a scan probe, it records the RSSI reading of the frame. The client will send this measurement back to the AP the next time it is allowed to access the medium. When the APs receive scan responses, they forward them to the controller. To reduce the probability that scan responses are lost, all APs that overhear the scan responses forward them to the controller.

There are four primary system parameters that determine the RSSI measurement collection: the interval and duration for periodic scanning, the interval over which link throughput is measured, and the threshold used to explicitly request a scan.

Link Throughput Measurement Interval and Threshold

An AP can explicitly request a scan when it detects a significant drop in throughput on any link. Table 4.1 shows the Allan deviation in the reported RSSI and packet reception rate across different time intervals for the least bursty and most bursty link in our testbed. The antennas are oriented based on the MaxCAP configuration. We find that the links exhibit varying behavior. For example, for a 100 ms window, the least bursty link only has a 15% deviation in reception rate while the most bursty link deviates by as much as 26%. Thus, it is difficult to select a threshold that works for all links and maintaining per-link thresholds will lead to a significant overhead. This led us to an alternative approach, where the detection thresholds are set to a conservative value and we rely on fixed-cost periodic scanning to prevent the system from operating in a suboptimal state. We measure link throughput over 100 ms intervals (which span five 20 ms time slots) and set the detection threshold for explicit rescanning to be 50% of the reception rate during the first window after a scan.

Interval	500 ms	1 s	5 s	30 s	10 m
Mean Thp (Mbps)	13.9	19.3	19.5	18.3	17
Std. Dev. (Mbps)	0.2	0.3	0.7	0.8	2.4
Overhead (%)	48	24	4.8	0.8	0.04

Table 4.2: Effect of scanning intervals on link throughputs

Periodic Scan Interval and Duration

The periodic scanning should be frequent enough to maintain an updated conflict graph and should not significantly impact the network performance. Table 4.1 shows that the deviations of the RSSI measurements are stable even for small time intervals, e.g., even for the most bursty link it is only about 2.2 dB for 5 ms. The RSSI measurements are much more stable than reception rate because they are obtained when all other APs are silenced while throughput measurements are obtained when multiple APs are active. This means that a 4 ms scan duration will accommodate most short-lived changes in RSSI readings.

To determine the scanning interval, we evaluate the overhead of scanning on the link throughputs. Table 4.2 shows the mean link throughput and corresponding scanning overhead for different scanning interval values. We see that when the scanning interval is too small, e.g., 500 ms, the overhead of the protocol is too high, reducing the link throughputs; while if the scanning interval is too large, e.g., 10 min, the system is unable to escape suboptimal states. The table shows that 5 seconds is a reasonable tradeoff, incurring less than 5% overhead while providing up-to-date RSSI information.

This system design is optimized for nomadic clients, i.e., clients that may appear at multiple locations across the network (laptops) but are used in stationary positions. In Chapter 4.4 we show that DIRC can also adapt to scenarios where users are walking with their devices, i.e., DIRC can handle scenarios where a client constantly moves. However. we decided not to consider a highly mobile scenario, e.g. vehicular networks, since these mobile applications are mostly latency-bound instead of capacity-bound, a case that may not be best addressed by DIRC.

4.3.2 Two-Phase Timeslot Protocol

DIRC relies on the intelligent scheduling of transmissions across the network. A timeslot based MAC protocol is used to synchronize the downlink and uplink traffic. Additionally, DIRC's timeslot protocol also adapts to different traffic workloads instead of having a fixed time allocation for the uplink and downlink traffic.

All nodes in the network, including APs, clients, and the controller are time synchronized. On the clients and APs, hardware clock synchronization is implemented in the Atheros 5212 card through beacon messages (and when all the nodes are in adhoc mode). Since the driver has access to the hardware clock, wireless nodes can transmit at the time slot boundaries with microsecond accuracy. Notice that the default HZ (frequency of timer interrupts) in the linux kernel used on the Phocus array is set to 100, causing the interrupt to be every 10ms. We manually set the HZ to 1000, but the frequency of interrupts is very unstable, and we are only able to accurately im-

plement a minimum timeslot of 20ms (and 4ms if accuracy requirement is relatively low, as we did in the timeslot split). We borrowed many implementation ideas from the CARP project [13]. Our timeslot-based MAC operates in two phases; the first phase is dedicated to directional transmissions, while the second phase accommodates omni-directional nodes, including clients and legacy APs. We term these two phases of our protocol `dirc-tx` and `omni-tx`.

This is similar to the split of contention and contention-free period in 802.11 PCF, but the difference is that the scope of contention-free in our protocol includes all APs and users, while the scope in 802.11 PCF is specific to that particular AP and its associated clients only.

During the `dirc-tx` phase, the centralized controller schedules AP-to-client transmissions, with carrier sensing disabled at the APs. This is done by raising both the receiver and transmitter CCA thresholds, allowing the APs to send frames even in the presence of interference. Due to some implementation errors in disabling carrier sensing in the Atheros HAL, we have to disable the receiver on the APs as well. Furthermore, the inter-frame spacing (IFS) is set to the minimum possible value since it is not needed in a timeslot schedule. We also disable link-layer ACKs since they introduce mixed traffic from both directions (and the receivers are disalbed so no ACKs can be received anyway). Then to handle frame losses, similar to previous work with a timeslot-based MAC [90], our protocol uses block link-layer ACKs, i.e., a bitmap that encodes all frames received in the previous time slot, transmitted during the `omni-tx` phase. These ACKs are particularly important to achieve good TCP performance.

Note that though individual channels are symmetric due to channel reciprocity, the SINR values are not, i.e., they do not apply if we switch the roles of all transmitters and receivers. For this reason, the `omni-tx` phase uses the traditional 802.11 protocol, with all nodes, including APs, using omni-directional antennas.

It is certainly possible to leverage directional antennas for client-to-AP communication, but it is more complex and will have more overhead, because it is harder to schedule the client traffic (since they are not on a wired network).

By default, our system operates on 16 ms `dirc-tx` phases and 4 ms `omni-tx` phases. There are two reasons for this choice: 1) the majority of the traffic in enterprise WLANs tends to be downlink, from the APs to the clients, thus the default size of the `dirc-tx` phase is set larger than the default size of `omni-tx` phase; and 2) 4 ms is the minimum slot time that can be scheduled reliably in our implementation. A fixed time allocation for directional and omni-directional transmissions, however, severely limits the types of workload that can be accommodated by the network. As a result, DIRC uses measurements to identify the right time division between the two phases. The controller periodically collects information about the queue length on clients and APs. It then calculates the client-AP backlog ratio and uses this to dynamically adjust the slot time sizes. For simplicity, our controller picks one of the three up-down ratios: 16-4 ms, 10-10 ms, and 4-16 ms. Starting from a split of 16-4 ms (down-up), the network will switch to 10-10 ms if the ratio drops below 1, and 4-16 ms if the ratio drops below 0.25. The same holds in the reverse direction. The slot time size are adjusted at most once every second to avoid thrashing, and to reduce the probability of inconsistencies among APs and clients. Note that as mentioned before, a more flexible solution that does not require splitting the timeslot is presented and used in Speed (Chapter 5).

4.3.3 Frame Transmission and Reception

The AP transmits frames destined to a particular client as follows. First, the AP sends a *request-to-send* to the centralized controller over the wired Ethernet connection. This request-to-send includes the AP and client identifiers, and the queue length on this AP; the controller discards additional requests containing the same AP and client identifiers.

The controller collects all of the requests-to-send and calculates a time slot assignment for the next `dirc-tx` time slot using Algorithm 3. The controller sends a *clear-to-send* message (over the wired Ethernet) to each AP that it has scheduled for the upcoming time slot. The clear-to-send message tells the AP which client it should transmit to next, what direction and what data rate to use when transmitting its frames. We chose to serve a single client during the entire time slot, since serving multiple clients will require a change in the antenna orientation. This may force clients to wait a time proportional to the number of clients associated with the same AP. We could reduce this time by (i) using finer granularity time slots (4 ms was the minimum time slot we could implement in our testbed), or (ii) allowing multiple clients to be served using the same antenna orientation (either because they are in the same sector, or happen to receive adequately strong signals under the same antenna configuration). Note that if no AP is active, and there is a change in the time slot split between directional and omni-directional traffic, the controller will assign the APs to send the new split information omni-directionally. This split information could be easily integrated in the Beacon frames.

When the `dirc-tx` phase arrives, each AP checks to see if it is clear-to-send. If not, the AP waits until the next time slot and checks again. If the AP is scheduled to send, it un-buffers the frames to that client, sets the direction, disables CCA, and transmits frames until the end of the `dirc-tx` phase. The AP appends a trailer to each frame that includes a frame identifier for a block ACK, and the split for the next time slot. Each AP buffers all unacknowledged frames. The AP estimates the transmission time for each frame that it transmits (or retransmits), so it can make sure that the total transmission time for the frames put on the hardware queue does not exceed the time slot duration.

During the `dirc-tx` phase, the clients receive frames from the APs and record all the frame identifiers in a bitmap. Then, in the `omni-tx` phase, each client first transmits a block ACK back to the AP. The use of block ACKs can result in out-of-order frames which significantly reduces the performance of TCP. To address this, frame reordering is implemented in the driver.

After the client transmits its block ACKs, it transmits any other frame. The queue length on that client is piggybacked in each frame, and forwarded to the controller by the AP. During this `omni-tx` phase, the APs all set their antennas to omni-directional mode, and start receiving frames from clients. When a block ACK is received at the AP, the corresponding frame is removed from the receive buffer. Note that in this phase, the wireless nodes operate using CSMA, e.g., link-layer ACKs are enabled and the default Madwifi rate adaptation algorithm is used.

4.3.4 Association

In Speed (Chapter 5), we show that the client-AP assocation problem in directional antenna systems is quite different from that in omni-directional antenna systems: it makes sense for clients to associate with one of the multiple APs even operating in the same channel. In DIRC

system, however, we do not consider the AP-client association problem, because the benefits of the flexible directional association with only directional APs are not very strong (as shown in Chapter 7, no client can gain 15dB separation from both APs). In this section, we discuss the association process after the client has decided to associate with an AP.

It is important for DIRC to quickly bootstrap a new client into the network and update the conflict graph while minimizing the overhead on the system. When a new client joins the network by sending an association request to the AP, the controller needs to generate the conflict graph related to the new client. The number of probes needed to generate the conflict graph are $O(N \times K_{AP})$ and may result in a high overhead when the association and disassociation rate is high. Thus, in our protocol, we take the lazy approach where measurements are delayed until the next scanning period. The lazy approach incurs minimum overhead for client association, but it has the downside that the conflict graph will be incomplete for this client. The conflict graph, however, must be present to determine whether two links will interfere with each other or not. The controller will not schedule transmissions when the part of the conflict graph for that transmission is incomplete. Thus, when the client first associates with an AP, the AP is not able to initiate frame transmissions to the client until the next scanning period. In practice, we do not expect this delay to be a problem because the client typically associates with an AP upon boot time. The inherent delays in this process, e.g., the clock synchronization in wireless hardware can take several seconds, is likely to mask the controller delays. Certain data traffic that may follow the association, such as DHCP requests, need to be delayed until after synchronization and scanning is finished. Note that if the RSSI information of the associating client is available on the controller (e.g., the client deassociates with one AP and associates with another AP), the controller can schedule it immediately.

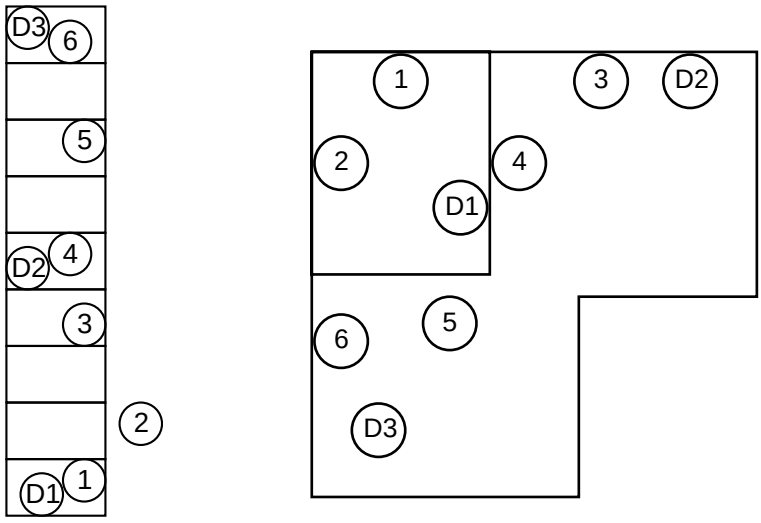
4.4 Evaluation

In this section, we evaluate DIRC's centralized scheduling algorithm and overall performance using a working implementation of DIRC in two real-world settings, Campus 1 and Lab 1 scenarios. We present three key results:

- DIRC's centralized algorithm can achieve close-to-optimal transmission scheduling and antenna orientations with very low overhead.
- The end-to-end DIRC implementation works well in practice. In a network with three directional APs, UDP performance improves by 65% over the MaxSNR approach (100% over using only omni-directional antennas), and TCP performance improves by 40% over the MaxSNR approach (42% over omni-directional antennas).
- DIRC can handle node mobility and dynamic traffic patterns.

4.4.1 Experimental Setup

Chapter 2 briefly discusses the experimental setup and how we collect measurements. Here we describe the two scenarios of Campus 1 and Lab 1 in more details. Notice that the measurements taken in these two scenarios are very different from other testbeds in that we actually measure



(a) Campus 1 (b) Lab 1

Figure 4.4: Experimental map for evaluating DIRC

Scenario	**# APs**	**# clients**	**AP ant.**	**client ant.**
Campus 1	3	6	45°,omni	omni
Lab 1	3	6	45°,omni	omni

Table 4.3: Experimental setup for evaluating DIRC

both the signal strength and the network capacity. While in other scenarios, we only measure signal strength and rely on the SINR model to compute the network capacity.

Our experimental testbed consists of three directional antennas (Fidelity Comtech Phocus Array [3]) and six omni-directional nodes (Linux desktops and/or laptops). The phased array antennas have eight antenna elements and the antenna pattern has a beamwidth of 45°. The picture of the antenna and the antenna patterns are shown in Figure 2.1(h)-(j). The antenna can be electronically steered at angular displacement of 22.5°, providing a total of 16 directional antenna states and one omni-directional antenna state. Since the goal of the protocol is to improve spatial reuse instead of extending communication range, the directional patterns we picked have similar antenna gain to that of the omni-directional pattern (only around 2 dB difference), and the side lobes of the directional patterns are very small.

We conducted our experiments in two physical testbeds: Campus 1 and Lab 1 scenarios. The maps and the locations of the nodes are shown in Figure 4.4. Nodes $D1 - D3$ are the directional antennas, and the omni-directional nodes are labeled $1 - 6$. Since the experiments use the 2.4 GHz band, they share the channel with other 802.11 devices. Table 4.3 summarizes the experimental setup for both scenarios.

Our evaluation consists of two parts. In the first part, we use collected measurements to evaluate the accuracy of the SINR model and the emulated performance of DIRC's core algorithm. The second part reports on the working end-to-end implementation of DIRC, with all associated

overheads.

4.4.2 Direction Selection and MAC Strategies

In this subsection, we first evaluate the effectiveness of the SINR based model to predict link throughput. We then evaluate various direction selection and MAC strategies using offline measurements, including the performance of DIRC's centralized algorithm for orienting the antennas and scheduling transmissions as presented in Chapter 4.2. In these experiments, we collect measurements first and then apply various algorithms on the collected data.

The data measurements are collected as follows:

- The APs broadcast UDP traffic, and all clients record the RSSI and link throughputs from the APs. This approach helps to reduce the number of measurements that need to be collected.
- To evaluate the effect of CSMA, we carry out one set of experiments where all three APs transmit using the default CCA threshold across all possible antenna orientations (i.e., $17^3 = 4913$ measurements).
- To determine the maximum achievable performance, we run a set of experiments to measure capacity under all possible schedules and antenna orientations (i.e., $\binom{3}{1}17 + \binom{3}{2}17^2 + \binom{3}{3}17^3 = 5831$ measurements). Here, we disable CCA on the APs to ensure that they all transmit simultaneously. Note that the maximum capacity results are drawn out of this particular experiment.
- We run the above-mentioned experiments twice for each of two data rates: 36 and 54 Mbps. These data rates are chosen because the AP-client distance is relatively short in our testbed.

In summary, in this offline experiment, we collect a total of $(5831 + 4913) * 2 = 21488$ measurements, and each individual measurement runs for 1 second. The whole experiment spans several hours, and the environment may change during the experiment. Since we apply the various algorithms on the same data set, all tested algorithms are consistently exposed to the same changes in the environment.

Also, we use the following technique to make sure that all three APs are transmitting simultaneously:

- We first send the *start-of-experiment* commands to all three APs.
- Then wait for the APs to send back acknowledgements.
- If not all acknowledgements are collected within some amount of time t, we abort and restart the experiment
- If all acknowledgements are rececived within t, indicating that all APs must have started the experiment at t.

In calculating the link throughput, we do not consider frames that are received within the first t time, and within the last t time, as these transmissions may not be synchronized. As mentioned above, all other frame receptions are from concurrent transmissions and should be taken into consideration. To change the antenna orientation, we used the *phasinfo* command provided by the vendor. We find that this command may fail sometimes for unknown reason, and requires a reload of the the *phasctrl* daemon to fix.

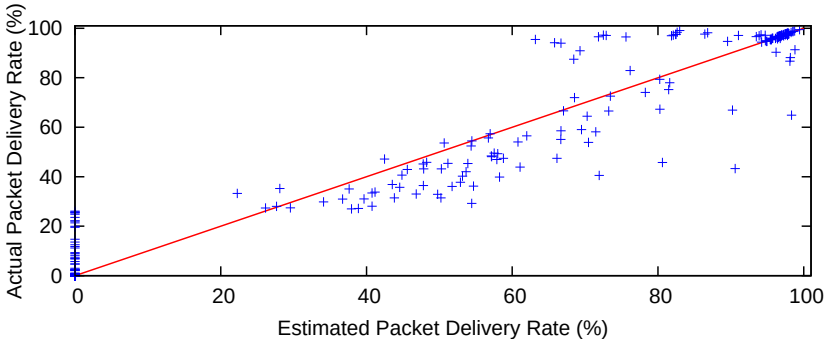
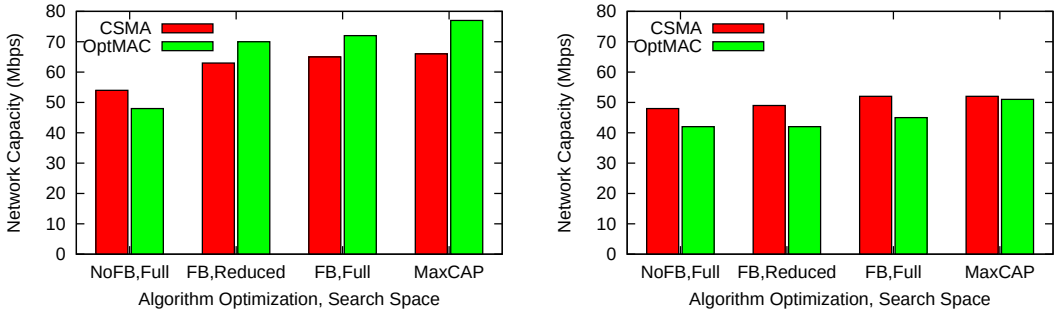


Figure 4.5: Estimated vs. actual packet delivery rate



(a) Campus 1

(b) Lab 1

Figure 4.6: Performance of the DIRC's scheduling and direction selection algorithm

In Chapter 3, we already show that the SINR model can predict packet loss rate reasonably well (Figure 3.2). Now let us look at how effective is to use the SINR model to predict link throughput. In order to achieve this, we calculate the correlation coefficient of the estimated throughput using the SINR model and the actual measured throughput. Figure 4.5 shows a scatter plot for a single client. The coefficient across all the nodes in the testbed is 0.98 which is sufficient for DIRC to pick the right schedules.

Client to AP association is not addressed by DIRC (but by Speed as presented in Chapter 5). To evaluate overall network capacity, however, we associate clients to the closest AP after the initial scan. We do that by looking through all possible associations of 3 clients (out of the 6 client locations in our testbed) with the 3 directional APs. Next, we evaluate the network capacity for a number of possible client-AP associations. We derive all possible sets of three clients out of the possible six locations ($\binom{6}{3} = 20$). For each three-client configuration, we test all possible client-to-AP associations, where each AP transmits to one client. For each configuration, we identify the configuration that leads to the highest overall capacity and use this as the most reasonable association pattern.

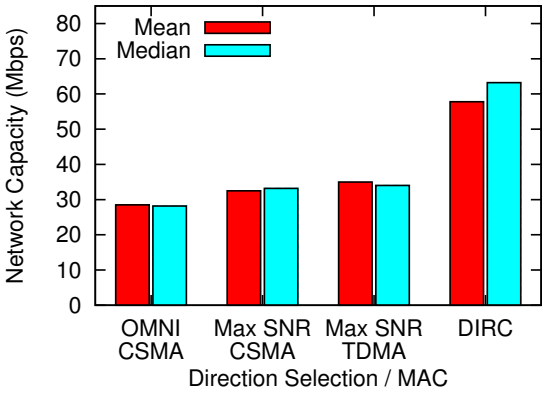
Before presenting the results, we should note that since we use broadcast frames to collect the measurements, our data does not incorporate the effect of rate adaptation. Thus, we cannot directly evaluate any interactions between the simple heuristics and rate adaptation. Therefore, in our evaluation, we emulate optimal rate adaptation for these heuristics, picking the best data rate for each individual link that achieves the max throughput. This provides an upper bound on the performance of these competing designs. Note that this does not apply to DIRC's algorithm since DIRC selects the data rate for each individual link as part of the algorithm.

Figure 4.6 compares the performance of three variants of DIRC's centralized algorithm to the maximum achievable capacity. *Reduced*, in Figure 4.6, refers to the scheduling heuristic for the reduction of the search space, while *Full* refers to the full scheduling algorithm. *FB* and *No FB* refers to the use of feedback to avoid problems with SINR-based throughput prediction. MaxCap is the maximum achievable capacity. DIRC's centralized algorithm is shown as FB/Reduced. The results show 1) the effectiveness of the SINR model in picking schedules and orientations, 2) the effectiveness of the feedback mechanism (FB), and 3) how the scheduling heuristics affect the network capacity and computation cost.

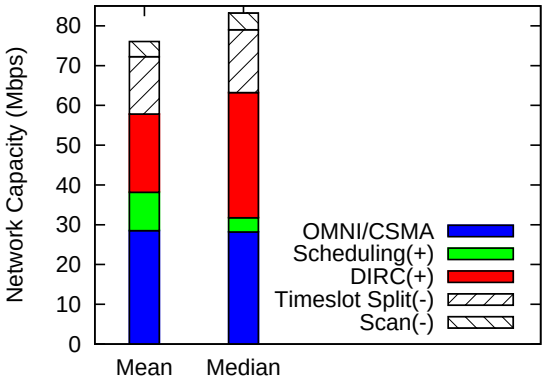
The results show that directly using the SINR model (No FB/Full) to pick the schedules often performs much worse than Max Cap. After applying the feedback mechanism that improves on the SINR model predictions (FB/Full), the performance is very close to that of Max Cap. DIRC's algorithm (FB/Reduced) leads to a slight reduction in network capacity, but its heuristics are able to reduce the number of configurations searched from 4912 down to 283 and 255 in two scenarios—a factor of 17X improvement.

4.4.3 End-to-End Performance

In this section, we evaluate the protocol and implementation of DIRC. We examine how much overhead is incurred, where the gain comes from, how well it interacts with transport protocols (especially TCP), and how well it can deal with environment dynamics and node movement. In this set of experiments, we deployed DIRC in the Campus 1 scenario. Since this set of experiments evaluates the end-to-end protocol performance that depends on the client responses, (e.g.,



(a) UDP Performance



(b) Overhead and gain for UDP

Figure 4.7: UDP performance of DIRC

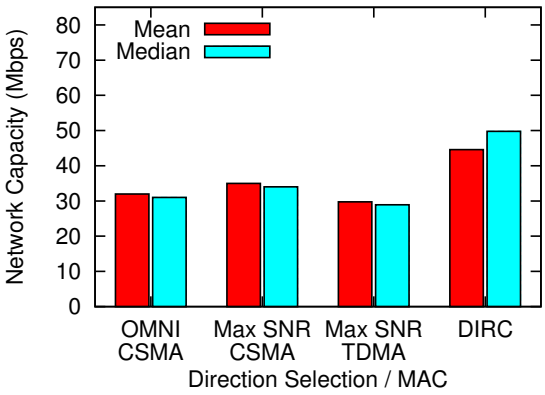


Figure 4.8: TCP performance of DIRC

block ACKs, scanning responses, TCP ACKs), we must pick a static AP-client association. For simplicity, we associate clients with the closest AP: $1, 2 \rightarrow D1$; $3, 4 \rightarrow D2$; $5, 6 \rightarrow D3$. We measure UDP and TCP throughput using the standard *iperf* utility. The packet generation rate for UDP is set to 30 Mbps (TCP manages the rate itself).

Figure 4.7 and Figure 4.8 shows the UDP and TCP performance of DIRC. In these experiments, the directional APs run *iperf* to each of their clients for 10 seconds, and we repeat this end-to-end measurement for each of the AP's clients (total of eight experiments). We report mean and median capacity. We also ran experiments for a longer time and the results are similar.

Although the UDP packet generation rate is 30 Mbps, the maximum actual throughput that DIRC can achieve is only 27 Mbps. This is because by default DIRC reserves 20% of the airtime for client transmissions, though it recovers some of that lost throughput by reducing the inter-frame spacing (IFS); as the maximum effective throughput for 54 Mbps is approximately 32 Mbps, the maximum throughput DIRC can achieve is approximately 27 Mbps. Consequently, the maximum network capacity for 3 directional transmitters is approximately 81 Mbps. Our

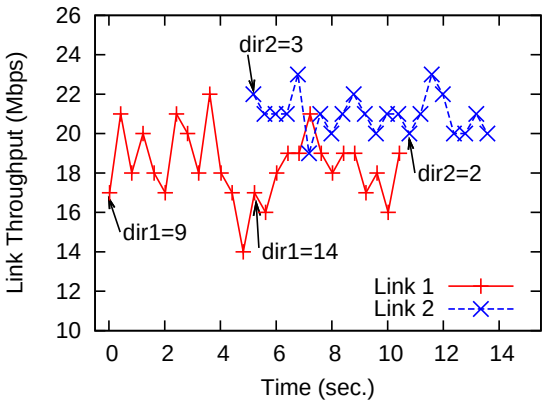
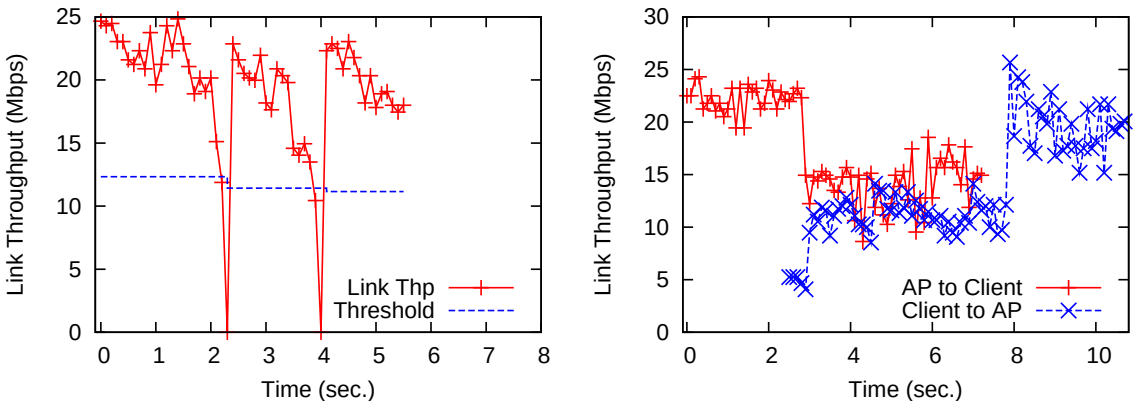


Figure 4.9: Protocol behavior during initialization

results show that the median UDP capacity of DIRC is about 76% of this upper bound, or a median concurrency level of 2.3. Note that CSMA MAC does not have the 20% loss from airtime reserved for the client (nor the gain from reduced IFS), allowing it to achieve close to 30 Mbps (out of 90 Mbps). Figure 4.7(b) shows the breakdown of DIRC's gain and overhead, where the solid part shows the gain and the patterned part shows the overhead. Note that some of these numbers are estimated, not directly measured. For example, the timeslot split is 16 vs. 4 ms, and the 4 ms *omni-tx* slot is wasted for UDP traffic, incurring a 20% loss. Then we use the actual network capacity to estimate the loss in Mbps from timeslot split (which is 1/4 of the capacity in Mbps). It shows that much of the benefits come from directionality of the antennas, and the major overhead for UDP traffic is the reserved 4 ms omni-tx phase. The TCP performance of DIRC is approximately 45 Mbps, which is a 40% improvement over max SNR and 42% over the default omni-directional antennas. The improvement is not as high as UDP, which can be due to 1) the interactions between timeslot and TCP, and 2) the split between two phases is not flexible enough.

Next, we study the performance of the protocol itself. We pick two links from the campus testbed: link 1 from $D1$ to client 2 and link 2 from $D2$ to client 3. We run *iperf* UDP tests with a packet rate of 30 Mbps for 10 seconds on both APs. We present the throughput of each link and the orientation of the directional APs in Figure 4.9. During $t = 0$ to 5, link 1 is the only transmitter and uses direction 9, which is the max SNR direction. At $t = 5$, scanning is carried out and link 1 suffers a drop in throughput due to the measurement collection. Since there are two transmissions now, the controller decides that two APs can transmit simultaneously, using directions 14 and 3. At $t = 10$, link 1 stops transmitting, and since link 2 is now the only transmitter, it can use MaxSNR direction (i.e., 2).

Figure 4.10(a) shows the protocol behavior when we use a laptop as the client and move from location 4 to location 2 in the campus setup, while associated with $D2$. Note that between any two scanning events, the movement of this laptop will only affect its own link throughput, but not the throughputs of other transmissions. Furthermore, for multiple mobile users, one rescanning will update the RSSI information for all mobile devices. Thus, we only enable one transmission in this experiment to simplify our presentation. We run the UDP *iperf* test from $D2$ to the client



(a) Handling node mobility

(b) Handling change in traffic pattern

Figure 4.10: Handling dynamics: node mobility and change in traffic patterns.

at the 30 Mbps data transfer rate. The detection threshold for re-scanning is set to 50% of the link throughput of the first window after each scan. At time $t = 0$, the client is at location 4, as it moves towards location 2, the link throughput drops below the detection threshold and triggers a re-scanning at $t = 2.3$; this process is repeated and the re-scanning is triggered again at $t = 4$; finally it stabilizes at around $t = 4.5$. The drops in throughput at $t = 2.3$ and $t = 4.0$ are due to scanning.

Figure 4.10(b) shows how DIRC can respond to changes in traffic patterns by adjusting the amount of airtime allocated to APs versus clients. In this experiment, we let client 4 associate with AP $D2$ and run UDP *iperf* on both nodes to each other (bidirectional UDP), and we do not enable other transmissions. We let the second data transfer start several seconds later than the first data transfer. The figure shows the link throughputs for both directions over time. At time $t = 0$, only the AP is transmitting and the scheduler allocates most of the airtime to the AP. At $t = 2.3$, the client starts transmitting. The queue length on the AP reaches around 30, and on the client it reaches around 54; these lengths exceed DIRC's 1/1 threshold, causing the controller to split the airtime equally between the AP and client at $t = 2.8$. At $t = 7.2$, the transfer from AP to client ends, and at $t = 7.8$, the controller re-allocates the airtime again, but favoring the client.

4.5 Related Work

4.5.1 Outdoor Directional Antenna Systems

We first examine related work in outdoor directional antennas. Existing work on directional antennas mostly assumes outdoor applications, where the dominating line-of-sight (LOS) path should be used for communication. This work includes MAC layer protocols for directional antennas [18, 33, 34, 65, 66, 91, 119, 122, 127], using directional antennas in vehicular networks [87], and using directional antennas to form long distance Wi-Fi links [31, 90, 97]. In Chapter 1 and Chapter 2, we show that because of the unique characteristics of rich scattering in indoors environments, these outdoor solutions do not perform well indoors. In particular, past

Protocol	Scenario	Ant. Conf.	MAC	Orientation	Goal
MobiSteer [87], etc.	outdoor	*omn-tx,dir-rx*	distributed	MaxSNR	connectivity
WildNet [90], etc.	outdoor	*dir-tx,dir-rx*	distributed	MaxSNR	comm. range
DMAC [65], etc.	outdoor	*dir-tx,dir-rx*	distributed	MaxSNR	spatial reuse
Midas [12]	indoor	*omn-tx,dir-rx*	distributed	MaxSNR	signal strength
DIRC	indoor	*dir-tx,omn-rx*	centralized	MaxCAP heuristics	spatial reuse
Speed	indoor	*dir-tx,dir-rx*	distributed	Greedy MaxCAP	spatial reuse

Table 4.4: Summary of directional antenna systems in various dimensions

schemes largely orient antennas by maximizing the signal strength (MaxSNR), which does not apply indoors.

4.5.2 Antenna Pattern Adaptation

The authors in [68] propose a very interesting technique to estimate the multipath channel between a phased array antenna and an off-the-shelf omni-directional client. Using this estimation, the directional APs can form adaptive antenna patterns that reinforce the signal strength at the recevier. This work is orthogonal to our work, and can be integrated into our system (Chapter 8).

One relevant effort [117] uses the capabilities of digital adaptive antennas to direct pattern nulls towards sources of interference. This is, in fact, similar to our idea of reducing interference to other transmissions. However, the application is still in outdoor scenarios, i.e., the direction of the null is the LOS direction from the sender towards the unintended receiver.

Antenna pattern adaptation has also been studied in the context of multicast [118]. In this work, the AP, after characterizing the locations of the clients, can use an adaptive antenna pattern that can multicast data towards multiple clients at the same time. To optimize multicast network capacity, the AP has to balance between servicing more clients at low data rates and servicing fewer clients at high data rates. In DIRC design, we only focus on unicast applications.

4.5.3 Other Directional Antenna Systems

Existing commercial directional antenna systems (Xirrus sectored WiFi array [8], Phocus array [3], Ruckus BeamFlex [5]) primarily focus on existing wireless networks today that are still noise dominated due to low level of contention. Thus these solutions are not suitable for our application scenarios which are mainly interference-dominated. We believe that given the popularity of wireless devices, more wireless networks will become interference dominated.

Midas [12] is a directional antenna systems that focuses on device orientation, e.g., how the orientations on the smartphones are changed and the implications for directional antennas on smartphones. Several observations made in this work are consistent with our findings, e.g., a small number of patch antennas suffice on size-limited devices. Directional antennas have also been used in other applications, such as localization [112].

Table 4.4 summarizes existing directional antenna systems and our two systems of DIRC and Speed along various dimensions, including application scenario, antenna configuration, antenna orientation algorithm, MAC protocl, and the optimization goal.

4.5.4 Scheduling Systems

There is also extensive work on scheduling in wireless networks using conflict graphs or physical layer interference information [17, 21, 22, 39, 41, 48, 52, 53, 67, 77, 78, 93, 98, 100, 101, 110, 126, 130, 131]. Some of this work is complementary to ours and might supplement or replace our basic scheduling heuristics in future versions of our system. Also, these scheduling systems look at various optimization goals (or utility functions), which can replace our optimization of maximizing spatial reuse if necessary, e.g., take into consideration of quality of service (QoS). Their integration, however, requires further study since our indoors scenario combined with directional antennas creates a much larger configuration state space.

4.5.5 MIMO Systems

In addition to spatial beamforming based on phased array antennas, MIMO beamforming can also be used to increase spatial reuse. Recent advances in multiuser-MIMO (MU-MIMO) techniques [15, 44] enable a single AP to serve multiple clients simultaneously by leveraging independent multipath channels at each client. MU-MIMO techniques require channel state feedback across all the clients in the network to characterize the multipath channel [1] and use advanced signal processing techniques like dirty paper coding and block diagonalization [35] to simultaneously transmit to all or a subset of clients.

4.6 Summary

In this chapter, we present the design, implementation, and evaluation of DIRC. DIRC is designed to maximize network capacity in indoor environments, especially enterprise networks where APs are directional, clients are omni-directional, and directional APs are centrally controlled.

Conventional wisdom has been that directional antennas are ineffective indoors, due to the complex scattering and multipath effects of the indoor RF environment. We demonstrate that this is not the case. DIRC is an indoor directional antenna system that is able to improve spatial reuse and system-wide network capacity in indoor spaces where wireless devices are densely deployed. In a nine node testbed, DIRC provides a 100% improvement for UDP and 42% improvement for TCP over omni-directional solutions, while being able to deal with changing workloads and environments. Notice that such conclusion is only validated with a specific experimental setup, as described in Chapter 1.3.1. The core of DIRC is an algorithm that identifies close-to-optimal antenna orientations with low overhead, and a MAC protocol that coordinates transmissions to maximize transmission concurrency. We implemented this design in a network using phased array directional antennas, and our experiments show that DIRC works well in practice.

Chapter 5

Speed: Directional APs and Clients for Future Wireless Networking

In the previous chapter, we present DIRC, a directional AP system designed for enterprise networks, and show that DIRC works well in practice. In this chapter, we will first take a step back and look at two main design dimensions of directional antenna systems: antenna configuration and antenna control. Then based on this exploration, we present the design and implementation of Speed [76], which is a directional AP and client system with distributed control.

Antenna configuration involves two questions. First, in Chapter 2, we show that putting additional directionality on clients can significantly improve spatial reuse. Since putting additional directionality on clients may involve much deployment efforts, the question is whether the two directional antenna configurations are fundamentally different, or can we achieve the same level of performance by putting narrower beam (that has even stronger directionality) antennas on the APs. In this chapter, we show that even with overall weaker directionality over the link, partitioning directionality across both APs and clients can provide a significant increase in network capacity, especially when clients are clustered. The second question is if we were to deploy directional antennas on the clients, what type of directional antennas should we use, and how to place these antennas on them? The biggest challenge is that wireless clients, e.g., laptops and smartphones, are usually size-limited and cost-sensitive. In this chapter, we show that due to the rich multipath in indoor environments, only a small number of narrow beam antennas are needed on directional clients, enabling a practical and less cumbersome directional antenna deployment on clients.

Antenna control involves three questions. First, the centralized solution in DIRC greatly limits its deployability; while the assumption of a wired connection and a centralized scheduler is reasonable in enterprise networks, it is rather unreasonable in other more general settings. Thus, a distributed solution is more desirable. The question is how to design this distributed protocol. A related question is how should individuals nodes decide what antenna orientation to use? We show in Chapter 2 that MaxSNR does not work well, but how about in systems with both directional APs and clients? Second, in DIRC, we assume that every client has determined which AP to associate with before running the DIRC protocol. The question is whether client-AP association problem in directional antenna systems is the same as that in the omni-directional antenna systems. If not, then how to associate clients to APs in directional systems? Third, the DIRC

protocol enables concurrent transmissions by disabling carrier sensing, and this implementation causes the senders unable to receive link-layer ACKs. Thus retransmission in DIRC is complicated and unpredictable, and out-of-order frames can happen. A more serious problem is that DIRC cannot co-exist with non-protocol-compliant 802.11 nodes. The question is whether there exists a better implementation that can achieve similar behavior and is more flexible at the same time?

Based on our exploration of the above questions, we describe the design and implementation of Speed. Speed is the first completely distributed directional antenna system that optimizes indoor wireless spatial reuse using directional APs and directional clients. While Speed can be applied to a variety of antenna configurations, in this chapter, we focus on one particular setup where Speed APs use phased-array antennas, and Speed clients are equipped with four patch antennas with an additional omni-directional antenna. In contrast to MIMO's multi-radio setup, Speed clients leverage antenna selection diversity and use a single radio. Speed has three components: 1) a timeslot based MAC protocol where each AP can reserve any timeslot that does not interfere with existing reservations on that timeslot, 2) an antenna orientation algorithm that each Speed AP runs to ensure non-interfering operation while transmitting directionally, and 3) a client-AP association algorithm that each Speed client runs periodically to determine which AP to associate with. Our experimental evaluation in two indoor testbeds shows that in practice Speed improves network capacity by up to 100% over existing solutions.

5.1 Antenna Configuration

5.1.1 Directional APs vs. Directional APs and Clients

In Chapter 2, we have shown that directional APs and clients can significantly improve the spatial reuse over the directional APs only solutions in two indoor testbeds (Campus 2 and Lab 2). Placing directional antennas on clients, however, involves much greater deployment efforts than replacing omni-directional APs with directional APs. Then the question is can we achieve the same level of performance with much narrower beam antennas on the APs (so that we do not have to deploy directional antennas on clients). In this experiment, we introduce another antenna configuration of *dir-tx,omn-rx* where APs use very narrow 16° antennas and clients use omni-directional antennas.

In order to compare the degree of directionality on two different antenna configurations (*dir-tx,dir-rx* and *dir-tx,omn-rx*), we propose a notion called directionality cost. The directionality cost for a directional antenna is defined as the number of antenna elements needed to implement that antenna in phased array: $\frac{360}{beamwidth}$. For example, 30° antennas need 12 antenna elements to implement, thus the directionality cost is 12. Also, the directionality cost for an omni-directional antenna is defined as 0. The directionality cost for a configuration is the sum of the cost on both APs and clients. For the definition of directionality cost, we ignore the difference in the number of APs (M) and the number of clients (N).

Using the directionality cost, we have the following three configurations with different costs:

- *dir-tx,dir-rx*, 35° APs, and effective 60° clients: the directionality cost is $360/35+360/60 = 16.3$

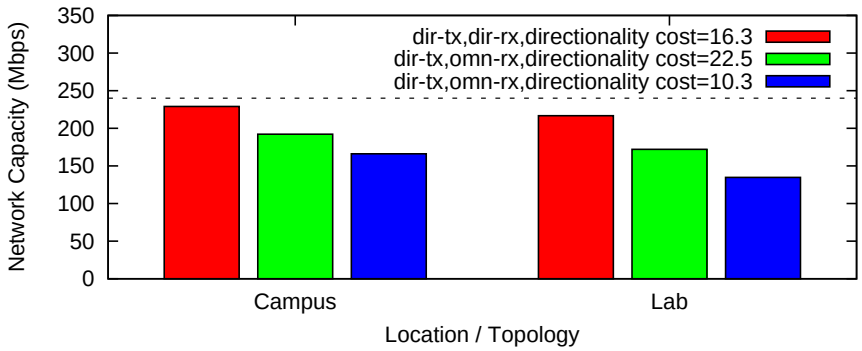


Figure 5.1: Network capacity with various directionality cost

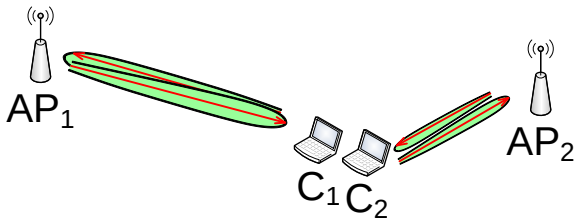


Figure 5.2: Fundamental difference between directional APs and clients and directional APs only

- *dir-tx,omn-rx*, 16° APs, and omni-directional clients: the directionality cost is $360/16 + 0 = 22.5$
- *dir-tx,omn-rx*, 35° APs, and omni-directional clients: the directionality cost is $360/35 + 0 = 10.3$.

Due to hardware limitations we are not able to compare different antenna configurations with the same directionality cost, but the three configurations listed here are enough to illustrate our point.

Figure 5.1 shows the network capacity for all three antenna configurations in both Campus 2 and Lab 2 scenarios. Comparing the two *dir-tx&omn-rx* configurations, we find that even when the total directionality cost is doubled from 10.3 (35° APs) to 22.5 (16° APs), the network capacity only improves by 15% and 27% for the campus and research lab testbeds respectively. However, distributing the directionality cost across the APs and the clients has a higher payoff. Surprisingly, even with a lower directionality cost of 16.3, the *dir-tx&dir-rx* configuration performs better than the *dir-tx&omn-rx* configuration with a directionality cost of 22.5. The network capacity improves by 37% and 60% over the *dir-tx&omn-rx* configuration with a cost of 16.3.

In order to gain insight of why this happens, Figure 5.2 shows an example to illustrates one possible reason. In this example, there are two senders transmit to two closely located clients (C_1 and C_2). When the clients use omni-directional antennas, the APs are unable to isolate the two clients (thus no spatial reuse) even when the APs use very narrow antennas. On the other hand, if the clients are equipped with directional antennas, spatial reuse becomes possible in this scenario. This is because if we look at one client C_1 and the interferer AP_2, they point at different

directions so the interference level will be low. This example illustrates the fudamental limitation of directional APs only configuration: it does not work well when the clients are clustered, and this clustering is common in several typical scenarios such as conference rooms.

5.1.2 How to Deploy Directional Antennas on APs?

As mentioned in Chapter 3, there are two types of directional antennas that are commonly used in indoor environments. Phased-array antennas use multiple antenna elements to form a directional antenna pattern, which can be oriented electronically in any direction. Their price however, is very high and their sizes are usually big. Patch antennas are smaller and cheaper, but they can only form a fixed sector of signal strength.

Since APs do not have strict size constraints, APs can use either sectorized patch antennas or phased-array antennas. In fact, in this chapter, we use both types on APs: in our measurements, we mechanically steer a patch antenna to emulate a directional AP; while in end-to-end protocol evaluation, we use phased array antennas on APs.

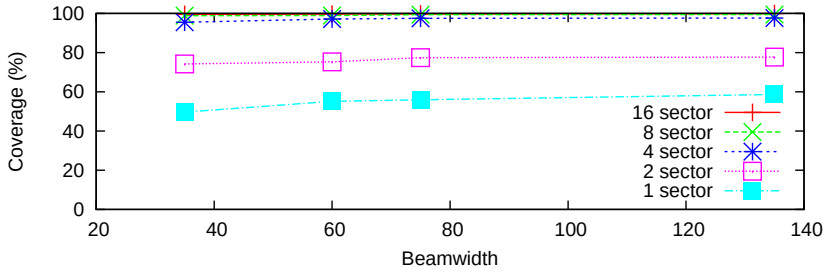
5.1.3 How to Deploy Directional Antennas on Clients?

We now consider what type of directional antennas to use on clients. Due to their low price and small size, patch antennas seem to be a reasonable choice on often size-limited wireless clients. The naive way to deploy patch antennas is to use antennas with a beamwidth that can cover the whole 360 degrees, e.g. four 90° antennas. In fact, to ensure full coverage, wider beams can also be used, e.g., 4 sectors of 120° antennas. Such provisioning ensures that the signal strength across the directional antennas will never be lower than the signal strength of a single omni-directional antenna configuration. This provisioning, however, turns out to be unnecessary in indoor environments. The rich scattering in indoor environments allows narrow beam antennas to provide similar coverage to that of wide beam antennas, even with a small number of sectors.

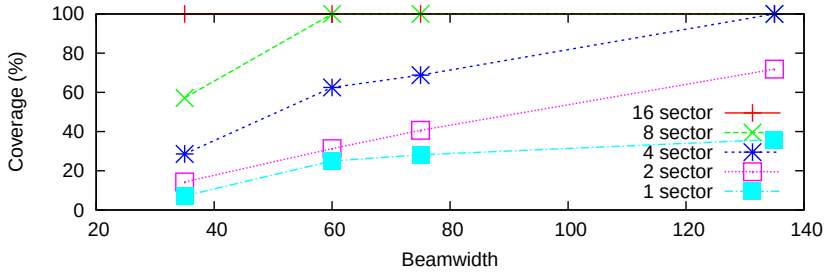
To demonstrate this result, we take measurements in three different locations: in a large outdoor space (Outdoor scenario), in a campus building (Campus 3), and in a research lab (Lab 3). Ideally we will deploy directional antennas on both senders and receivers, however, in order to obtain an extensive set of locations and physical orientations of the client (we obtained a total of 10624 data points), we deploy 6 omni-directional APs and 6 directional clients in the Campus 3 scenario, and 10 omni-directional APs and 13 directional clients in the Lab 3 scenario. For the directional clients, we use antennas with different beamwidths (35°, 60°, 75°, and 135°) and number of sectors (1 to 16). In each experiment, we measure the signal strength from all APs to all clients, across the 64 possible orientations of the client. The experimental maps are shown in Figure 5.8.

Figure 5.3 shows the coverage for the two indoor and one outdoor environments. We define coverage as the percentage of cases where the signal strength is stronger than that of omni-directional antennas. The result is averaged over all directional client/omni-directional AP pairs, and all 64 physical orientations of the client (thus a total of $(13*10+6*6)*64 = 10624$ location pairs).

Figure 5.3(a) shows that in indoor environments, even with only 4 sectors, 35° antennas can provide as good a coverage as wider beam antennas. As a comparison, Figure 5.3(b) shows

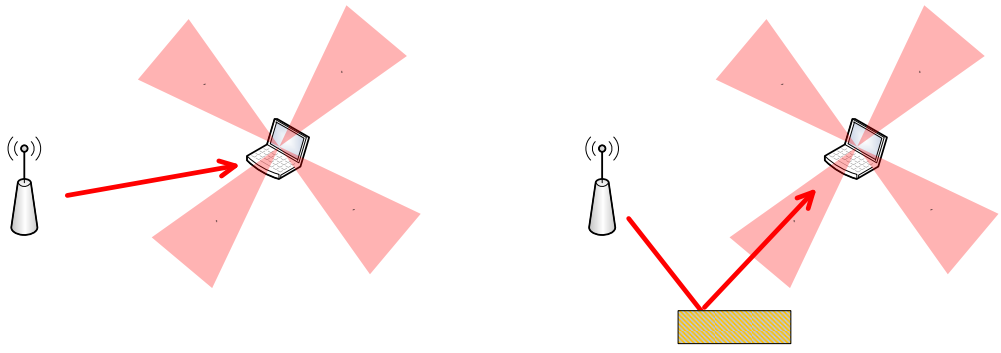


(a) Indoor



(b) Outdoor

Figure 5.3: Coverage of directional antennas



(a) Outdoor (b) Indoor

Figure 5.4: Fewer antenna sectors are needed to provide indoor coverage

that in the outdoor scenario, it requires 16 sectors of 35° antennas to provide good coverage. The reason for the difference is the rich scattering observed in indoor environments that provides multiple antenna orientations with comparable or higher signal strength than the omni-directional antenna. Figure 5.4 shows one example, when the client is equipped with 4 sector of narrow beam antennas, the AP may not be able to reach one of the sector in outdoor scenarios (Figure 5.4(a)), but with reflectors in indoor scenario, the AP may be able to reach one of the receive sectors through a reflector (Figure 5.4(b)).

Given the number of sectors, we should choose the minimum beam width that has good coverage (e.g., for four sectors, 35° antennas should be used) since this offers the stronger angular separation. For example assume a directional client is receiving a frame and an interfering signal. If the client is equiped with 35 degree sectors, it is likely to pick up less of the interfering signal than if it is using 135 degree antennas. If the client is equipped with four 35 degree sectors, the probability that the receiving sector will experience strong interference is 0.5 (i.e., the percentage of coverage with one sector as shown in Figure 5.3(a)). If the client is equipped with four 135 degree sectors, however, the probability of interference increases to 0.6.

Thus, we propose that directional clients that are primarily used in indoor environments should be equipped with multiple narrow beam antennas: 35° or 65° patch antennas, along with an omni-directional antenna to handle the scenario where none of the directional sectors can provide strong signal strength. Note that a typical 35° patch antenna is large in size, but antennas with a fan beam, e.g., 35° in one plane and 135° in the other plane, can be very small. In our measurement-based study, each client is equipped with two 35° fan beam antennas and two 65° patch antennas. This configuration will be practical for a laptop, but the engineering of new types of patch antennas and antenna placement is out of scope for this dissertation.

5.2 Antenna Control

By exploring antenna configuration, we have answered the two questions asked at the beginning of this chapter: we show that *dir-tx&dir-rx* antenna configuration can provide better performance than *dir-tx&omn-rx* even when using wider beamwidth antennas, and that there exists practical antenna placement on wireless clients. We now explore the design choices for antenna control for the *dir-tx&dir-rx* antenna configuration. Antenna control consists of three main sub problems: the design of an antenna orientation algorithm, an appropriate MAC protocol, and a mechanism for the association of directional clients to APs. We summarize the limitations of existing approaches and highlight the challenges involved in developing a solution.

5.2.1 Antenna Orientation and MAC Protocol

As presented in previous chapters, in outdoor systems, the simple strategy of orienting in the direction of maximum signal strength (MaxSNR) is best. Choosing antenna orientations for indoor directional antennas is harder because there are multiple paths between the AP and the client, and the optimal choice of antenna orientation depends on interfering transmissions. Below, we present the evaluation of two algorithms for *dir-tx,dir-rx* case: MaxSNR, an algorithm that chooses antenna orientation such that the AP-client signal strength is maximized, and Max-

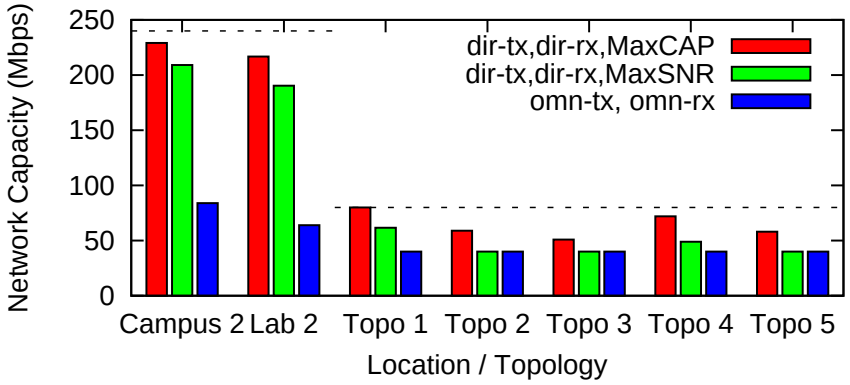


Figure 5.5: Network capacity with two orientation algorithms

CAP, an algorithm that picks the optimal antenna orientations by doing an exhaustive search over all antenna orientations.

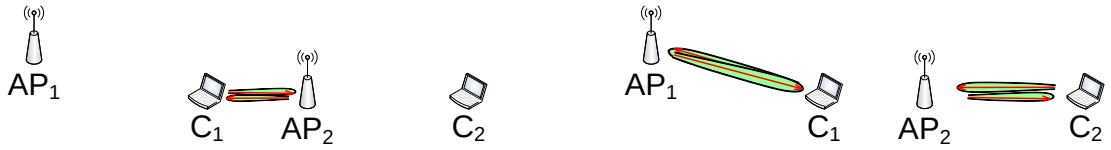
Figure 5.5 shows the network capacity from the two antenna orientation algorithms in several scenarios (Campus 2, Lab 2, and Topo 1-5). The Topo 1-5 are in a more controlled environment: a large room with tables and machines around, where we construct the different topologies of transmissions as shown in Figure 5.8(d)-(h), i.e., Topo 1–5. From Topo 1 to 2, the two transmissions are getting physically closer to each other. For Topo 3 to 5, we assume the two clients are co-located and we change the location of the APs. In Topo 3 and 4, the two APs are placed almost equidistant to the clients, and in Topo 5, one of the AP is a bit farther away from the clients than the other AP.

The figure shows that MaxCAP outperforms MaxSNR by about 15% in the Campus 2 and Lab 2 scenarios, and by much more, up to 47%, for the constructed topologies.

A closely related design decision is the choice of MAC protocol that coordinates the APs to choose antenna orientations for the transmissions and to identify non-interfering transmissions. DIRC (Chapter 4) has used a centralized scheduler-based MAC using MaxCAP. However, these designs are intrusive and require coordination among APs, making it unsuitable for some deployments, e.g., in wireless hotspots and neighborhood wireless networks. On the other hand, MaxSNR can be implemented in a distributed manner, but, as shown above, the performance is much worse. Thus, the challenge is to design a distributed antenna orientation algorithm and MAC protocol that can perform as well as its centralized counterpart.

5.2.2 Client-AP Association

In omni-directional antenna networks, the AP selection problem can be considered as a channel selection problem [16, 19, 25, 45, 94, 99, 102, 114, 120]. When the client chooses a particular channel, it always associates with the closest AP within that channel. For the *dir-tx&dir-rx* setup, even within the same channel, the directional client may have multiple APs that it can associate with due to the angular separation of the directional antennas, i.e., when C_1 is orienting towards AP_1, it is pointing away from AP_2 thus receiving less interference. Here, naively associating a client with the closest AP may be a suboptimal choice. One example is shown in Figure 5.6,



(a) MaxSNR assocation (b) More flexible association

Figure 5.6: Assocation in directional antenna systems

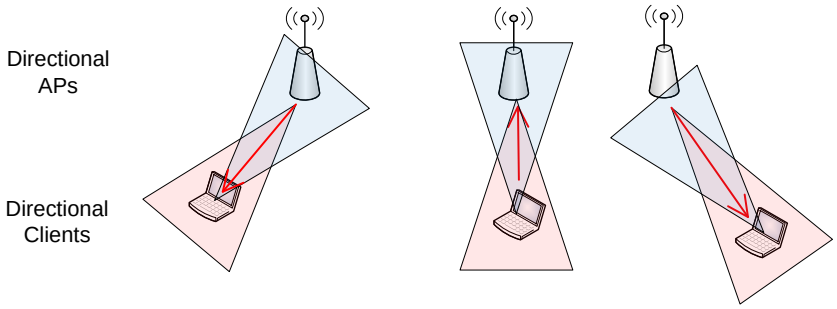
where C_1 is closer to AP_2 than AP_1. Thus it makes sense for C_1 to assciate with AP_2 most of the time (Figure 5.6(a)). However, if now C_2 also associates with AP_2, the link throughput from AP_2 to C_1 will be halved (AP_2 will spend half of the time servicing C_2). So if C_1 has a more flexible association strategy that allows it to associate with AP_1 which is idle then, two flows can occur simultanesouly (Figure 5.6(b)). Note that suppose all nodes are equipped with omni-directional antennas, it does not make sense for C_1 to associate with AP_1 because this transmission will always be interfered by AP_2. The key that enables this flexible association is the angular separation provided by the directional antennas. Another example is shown in Figure 5.8(f), Topo 4. Since both clients are co-located, they will always associate with the same AP with the naive association of closest AP. Again, a better choice is for each client to associate with a different AP to improve spatial reuse. Though associating the client with the most idle AP works in this scenario, it may be suboptimal in other scenarios. For example, consider Topo 4, but this time assume that there is a third AP AP_3 at the exact location of AP_1. In this case, the second client should still associate with AP_2, even if AP_2 has more traffic load than the newly added AP_3. This indicates that client-AP association needs to consider both the traffic conditions and available antenna orientations that minimize interference. The key challenge is to design an association mechanism that considers both factors.

5.3 Speed Design

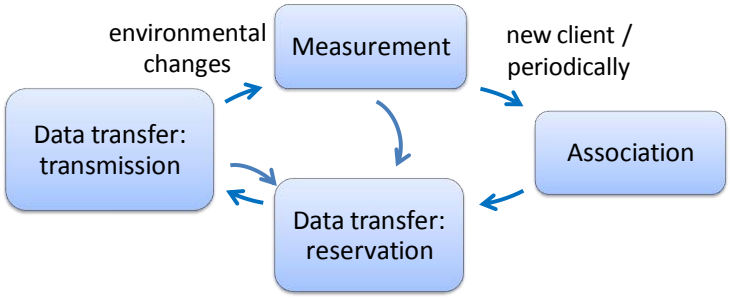
In this section, we present the design of Speed, a distributed directional antenna control system that optimizes spatial reuse. Speed's antenna setup is based on both APs and clients equipped with directional antennas. The antenna control system uses conflict graphs (Chapter 3) and traffic load estimates to overcome the unique challenges of antenna orientation, MAC protocol, and client-AP association posed by a network of directional APs and clients. Figure 5.7(a) shows Speed's system overview.

5.3.1 Assumptions

In Speed, we assume that APs and clients are equipped with a single radio that provides antenna selection diversity, which existing commodity Wi-Fi radios already support. This way, for each incoming frame, the RSSI readings on all antennas/sectors can be measured at the same time at the client.



(a) Speed Overview



(b) Speed Operation

Figure 5.7: Speed overview and operation

5.3.2 Overview

The Speed system operates in three phases: 1) measurement phase, 2) association phase, and 3) data transfer phase. Figure 5.7(b) shows these three phases, with two sub-phases in the data transfer phase.

The goal of the measurement phase is to construct the conflict graph at every node in the network. This is done by each AP sending probe messages across all its antenna orientations and clients recording the RSSI across all their antennas and reporting the results back to all APs. The measurement process is described in detail in Section 5.3.3.

In the association phase, the client uses the conflict graph obtained from the measurement phase and traffic load estimates embedded in the AP beacons to determine which AP to associate with. The association process is presented in Section 5.3.4.

Finally, APs use the antenna orientation algorithm and the supporting MAC protocol for the data transfer on the downlink. As mentioned in previous section, the distributed MaxSNR algorithm does not perform as well as the centralized MaxCAP algorithm. To manage this trade-off, Speed uses a new distributed algorithm which is a compromise between the MaxSNR and MaxCAP approaches. Speed's MAC protocol is based on timeslots (20 ms each with a beacon interval of 1 sec) and timeslot reservations. The idea is that a Speed AP can reserve a timeslot if it can find antenna orientations that do not interfere with existing reservations in that timeslot. This way, unlike MaxSNR, Speed does consider the orientations of other transmissions in the network to determine antenna orientations; and unlike MaxCAP, it does not rely on exhaustive search and can be implemented without requiring centralized coordination. Speed APs disable random backoff on their data traffic transmit queues [55, 109]. This way all the APs will be synchronized to transmit with each other after any data transmission (since there is no randomness), which effectively disables carrier sensing for data traffic. Speed AP rely on carrier sensing (with random backoff) only for other traffic such as timeslot reservations, traffic from clients to APs, traffic from external APs, etc. Thus, during each timeslot, APs concurrently transmit non-interfering data traffic to achieve spatial reuse. Other traffic, such as uplink traffic from clients to APs, traffic from other non-protocol compliant transmissions, and all management frames such as timeslot reservations, may be transmitted anytime, but these frames rely on CSMA mechanism to avoid collisions. The data transfer phase is presented in Section 5.3.5.

5.3.3 Measurement Collection

The goal of this phase is to collect the signal strength information from all APs to all clients with all possible antenna orientations on APs and clients, i.e., the complete table of $S(AP_i, C_j, K_{AP_i}, K_{C_j})$. As mentioned in Chapter 3, this table is needed to construct the conflict graph on each node. To obtain this information, each AP sends a number of frames using each of its orientation, and the clients record the RSSI readings of all received frames from each of its sectors. We assume the received signal strength at each receive sector can be measured simultaneously for any incoming frame, which is a valid assumption because existing techniques such as selection diversity already exploit this capability. Note that the measurement process in Speed is similar to that in the DIRC system in Chapter 4, but this process in Speed is fully distributed.

Measurements must be taken when a client joins the network, and in Speed, the association

process is initiated after the measurement process. The new client first sends a *request-to-scan* frame to the network, notifying all the APs of the new client. Upon receiving a frame, each AP can initiate its scanning process, which will take five timeslots, or 100ms. Thus, before the scanning process, each AP needs to reserve the next five timeslots by sending its *request-to-scan* frame. If a second AP also needs to initiate its scanning process, it can reserve the subsequent five timeslots by sending its *request-to-scan* frame in the last timeslot of the existing scanning process. In order to avoid frequent measurements, each AP postpones the scanning process if its last scanning finishes in less than one second. Note that these parameters may need to adapt to the number of neighboring APs of a network, which we left as future work, as the current parameter of one second can only accommodate 10 neighboring APs.

During scanning, the AP transmits a number of scanning probe frames (5ms for each orientation, thus 80ms or four timeslots for 16 orientations) to the network using the lowest data rate. The scanning probe includes the AP and the orientation used. All clients that receive these scanning probes record the mean RSSI readings, the AP identifier, the AP orientation,

5.3.4 Association

In Speed, clients determine which AP to associate with. The goal of the association phase is for the client to associate with the AP that can provide the highest estimated throughput. We assume that clients have already selected their channel of operation. To simplify the presentation, we ignore the AP backbone capacity, which is also an important metric in AP selection [88, 115] and assume that the wireless link is the bottleneck.

The estimated throughput is based on the conflict graph and the traffic load on each AP. The conflict graph is constructed based on the measurements collected. To estimate the traffic load at the client, each AP embeds its traffic load in its beacon messages, which includes the number of timeslots already allocated to clients associated with that AP $(t_1, ..., t_n)$, and the number of remaining idle timeslots t_{idle} on that AP.

The throughput estimation involves two parts:

1. the estimated maximum number of timeslots that can be allocated to the client, t_{exp}; and
2. the expected link throughput for each timeslot, rt_{exp}.

Then the estimated link throughput can be calculated as $t_{exp} * rt_{exp}$.

First, suppose the total number of timeslots between beacon intervals is t_{total}, where $t_{total} = \sum_i t_i + t_{idle}$. The client estimates the maximum number of timeslots t_{exp} that can be allocated to it if it associates with that AP as follows:

- If $t_{idle} \geq \max_i(t_i)$, then $t_{exp} = t_{idle}$,
- otherwise, $t_{exp} = \max_i(t_i) \times \frac{t_{total}}{\max_i(t_i)+\sum_i(t_i)}$.

The idea is that if the current AP is very idle, then the client can potentially use all the idle timeslots. And if the AP is relatively busy, i.e., the AP do not have enough idle timeslots to accommodate the request, the AP will assign the timeslots proportionally to the requests.

Second, the client estimates the link throughput rt_{exp} in each usable timeslot using the conflict graph. Algorithm 4 shows how a client C_j estimates link throughput if it associates with AP_i (several notations and functions used in the algorithms are defined in Chapter 3). The algorithm assumes that the traffic condition is stable from last beacon interval. The idea is to calculate the

link throughput from the best antenna orientations that can coexist with transmissions from other APs (lines 6–21). The client first calculates the link throughput if that link operates on its own, rts (line 4), which is used to calculate link throughput when there is no interference. Then for each antenna orientation combination (line 6), and for every other AP (line 8), it computes the expected link throughput with regard to that interfering AP, $currt$. This computation accounts for the probability of transmission from the interfering AP (lines 10–16) and the probability of that AP being idle (line 17). If the interfering AP is transmitting, then the link checks whether both links can co-exist, i.e., whether the sum of the link throughputs are larger than a particular threshold (line 14). Here $sinr\_to\_thp$ is defined in Chapter 3 which returns the link throughput for a particular SINR level. If both links can be active, then $currt$ is increased. If the interfering AP is idle, then the current link can be active and $currt$ is increased. If there are multiple APs, the expected throughput is determined by the strongest interfering AP (line 18). Finally, the expected throughput is chosen among all possible orientations (line 20). Then the expected link throughput from an AP is calculated as $t_{exp} * rt_{exp}$.

Since traffic from each client changes over time, to maximize performance, association should be determined periodically to best balance AP traffic. First, to avoid the client from jumping back and forth among APs, it only switches to a new AP when the expected improvement is above a threshold, whose optimal value depends on specific traffic patterns. We set this threshold to 30% in our experiments. Second, to avoid flash crowd, in which an AP becomes idle and becomes attractive to all the clients around, the reassociation decision should be probabilistic. Ideally, the probability should depend on the number of clients in the network, but in our system, we simply set the probability to 0.5.

5.3.5 Data Transfer

After association, the AP may send data frames to the client, using the distributed antenna orientation algorithm and MAC protocol described in this section. Note that the directional MAC protocol in indoor environments for data transfer has two functions: selecting the antenna orientations and identifying non-interfering transmissions.

The MAC protocol ensures non-interfering operation across directional transmissions by allowing new reservations only if they do not interfere with existing reservations. The process can be described in the following steps. First, each AP records all received *request-to-send* reservation frames. Second, when an AP has traffic to send to a client, it checks whether there exists any antenna orientation combination (Algorithm 5, line 2) such that this transmission will not interfere with existing reservation of the timeslot (line 4-9). If no such orientation exists, the AP cannot reserve the timeslot for the client ($rt_1 = 0$ in Algo 5). Otherwise, the AP prepares the *request-to-send* frame that includes the sender, receiver, the computed antenna orientations, and data rate. Note that unlike MaxSNR, Speed's Algorithm 5 does consider antenna orientations of other active transmissions, and unlike MaxCAP, it involves only local decisions that do not require a global exhaustive search.

The *request-to-send* frame is a management frame and is sent using CSMA/CA. Thus, each competing *request-to-send* frame has equal opportunity to access the channel. When the frame has been successfully transmitted, the reservation has been confirmed, which means that all the other APs can transmit only if they can do so without interfering. If an AP receives a *request-to-*

Algorithm 4: For association: C_j calculates expected throughput per timeslot from AP_i

Output: expected throughput per timeslot rt

$rt \leftarrow 0$
$rts \leftarrow 0$
foreach k_{AP_i}, k_{C_j} **do**
    $rts \leftarrow \max(sinr\_to\_thp(S(AP_i, C_j, k_{AP_i}, k_{C_j})), rts)$
end
foreach k_{AP_i}, k_{C_j} **do**
    $rtdir \leftarrow 54$
    foreach $AP_m \neq AP_i$ **do**
        $currt \leftarrow 0$
        foreach C_n *that associates with* AP_m **do**
            foreach k_{AP_m}, k_{C_n} **do**
                $sinr_1 \leftarrow PWS(i, j, m, n)$
                $sinr_2 \leftarrow PWS(m, n, i, j)$
                if $sinr\_to\_thp(sinr_1) + sinr\_to\_thp(sinr_2) > thresh$ **then**
                $currt \leftarrow currt + (sinr\_to\_thp(sinr_2)) * t_{C_n}/t_{total}$
            end
        end
        $currt \leftarrow currt + t_{idle}/t_{total} * rts$
        $rtdir \leftarrow \min(rtdir, currt)$
    end
    $rt \leftarrow \max(rt, rtdir)$
end
return rt

send from another AP while its own *request-to-send* is enqueued , then the AP will remove its own *request-to-send* from the transmit queue and will go back to the second step of Algorithm 5. Also, when a client receives a *request-to-send* intended for it, it will use the receive orientation specified in the *request-to-send* frame for reception. Note that since there may be multiple intended recipients of such *request-to-send* frames, there is no ACK generated. Instead, Speed relies on duplicate *request-to-send* frames and the use of low data rates.

A common technique to support concurrent transmissions is to disable physical layer carrier sensing at the APs (as in DIRC). However, as mentioned before, this will cause collisions with unscheduled uplink traffic (e.g., TCP ACKs) and other non-protocol compliant traffic. To mitigate this problem, downlink data frame transmissions are subject to two constraints. First, all transmissions occupy the same airtime independent of the data rate. And second, frame transmissions are synchronized. This maximizes spatial reuse achieved in a network with rate diversity. We satisfy these constraints by fragmenting and padding data frames to the airtime required for transmitting 1500 byte packets at 54 Mbps and disabling random backoff [55, 109]. The synchronized transmissions on the downlink lead to synchronized link-layer ACKs causing ACK collisions. However, surprisingly we observe that the synchronized ACKs were rarely corrupted.

Algorithm 5: For data transfer: AP_i determines orientations for C_j

Input: set of received *request-to-send* frames R, set of orientations on AP K_a, set of orientations on client K_c

Output: orientations used on AP dir_{AP_i}, on client dir_{C_j}, and rate rt_i; $rt_i = 0$ indicates no concurrent transmissions possible

$rt_i \leftarrow 0$; $dir_{AP_i} \leftarrow -1$; $dir_{C_j} \leftarrow -1$
foreach k_{AP_i}, k_{C_j} **do**
    $currt \leftarrow 54$
    foreach $(AP_m, C_n, k_{AP_m}, k_{C_n}, rt_m) \in R$ **do**
        $sinr_1 \leftarrow PWS(m, n, i, j)$
        $sinr_2 \leftarrow PWS(i, j, m, n)$
        if $sinr_1 > thresh(rt_m)$ **then** $currt = \min(currt, sinr\_to\_drate(sinr_2))$
        else $rt \leftarrow -1$; break
    end
    if $rt > rt_i$ **then**
        $rt_i \leftarrow currt$; $dir_{AP_i} \leftarrow k_{AP_i}$; $dir_{C_j} \leftarrow k_{C_j}$
    end
end
return $(rt_i, dir_{AP_i}, dir_{C_j})$

The same observation has been made by previous work [109]. All management frames on the downlink are enqueued in a separate hardware transmit queue where default CSMA is used. In the case of the uplink traffic, the AP may not be at the optimal receiving orientation. However, since these frames are transmitted with regular CSMA and are non-colliding (i.e., with no or low external interference), APs are still able to successfully decode these frames.

Data Rate Adaptation

As mentioned in Chapter 3, the SINR model can incorporate multiple data rates. Here we look at the support of data rate adaptation. Since the SINR level is independent of the data rate used on the interfering link, Speed APs can support existing rate adaptation schemes during a timeslot. In Speed, each AP picks the data rate to be used at the beginning of the timeslot by comparing the current SINR level to the minimum required SINR thresholds for the different rates. However, during the timeslot the selected rate may not be supported due to an inaccurate SINR model, uncalibrated RSSI levels, short-term fading, or mobility in the environment. In these cases, Speed APs can adapt their data rate without impacting other simultaneously transmitting links. While Speed can accommodate a wide range of rate adaptation algorithms, we use the simple auto rate fallback mechanism in our system.

Handling Dynamicity

In Chapter 4, we have a thorough discussion on handling dynamicity. The Speed protocol follows the same philosophy, and we summarize several key points. One potential problem of having

Scenario	# APs	# clients	AP ant.	client ant.
Campus 2	6	6	16,35°,omni	4×35° fan beam
Lab 2	6	6	16,35°,omni	4×35° fan beam, omni
Topo 1-5	2	2	35°,omni	4×35° fan beam, omni
Campus 3	6	6	all	omni
Lab 3	13	10	all	omni

Table 5.1: Experimental setup for evaluating Speed

strong directionality on both APs and clients is that performance can be much more sensitive to medium dynamics, i.e., current beams are blocked, or the client moves around. Dynamicity in these scenarios can be categorized as 1) changes that affect the current transmission, e.g., node mobility and environmental changes that happen in the sender and receiver beams, and 2) changes that do not directly degrade the performance of current transmission; instead, they make other antenna orientations more attractive. Since in Speed, the client can measure the signal strength of an incoming frame at different receive sectors simultaneously, both mobility and environmental changes may be detected at either the AP or the client. An AP can detect the performance degradation by observing a throughput drop due to frame losses or reduced data rates. A client can further observe changes in the AP's signal strength across different sectors.

5.4 Evaluation

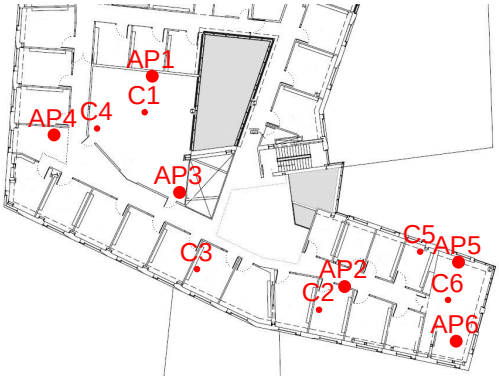
In this section, we evaluate the Speed antenna control system described in the previous section. Our evaluation has two parts: 1) a measurement-based evaluation, and 2) an end-to-end evaluation. The purpose of the measurement based evaluation is to perform controlled experiments and compare the performance of Speed's antenna control algorithm with other approaches across different network topologies. To better understand the measurement based evaluation of the different antenna control algorithms, we also use a heuristic, called the separation metric. We will present the details of different variants of the separation metric in Chapter 7. In this section, we use one variant of the separation metric (link pair metric) that captures the SINR values of the links. Finally, we present an end-to-end implementation and evaluation of Speed and address the challenges of implementing the system using commodity hardware.

5.4.1 Measurement-based Evaluation

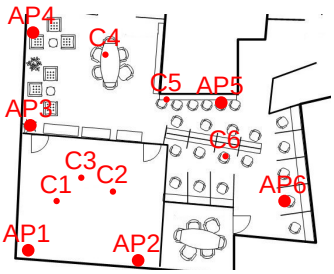
In this section, we present the measurement based evaluation.

Experimental Setup

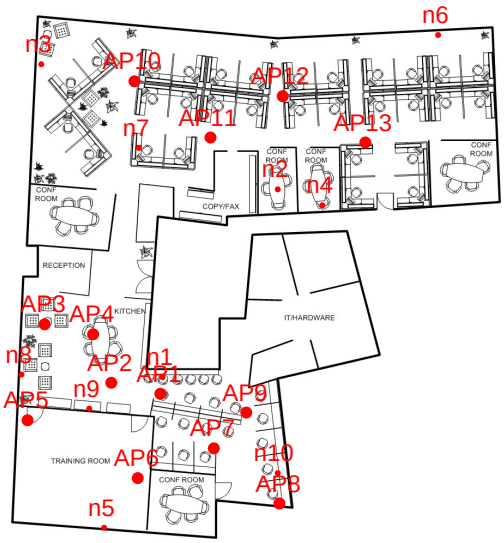
This experiment consists of two campus scenarios (Campus 2 and 3), two lab scenarios (Lab 2 and 3), and five typical topologies of two flows (Topo 1-5), i.e., Figure 5.8 and Table 5.1. The



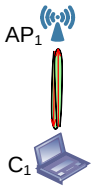
(a) Campus 2 and Campus 3

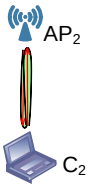


(b) Lab 2

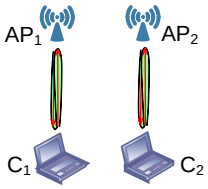


(c) Lab 3

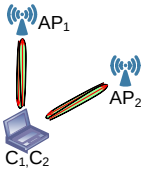




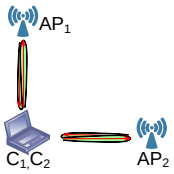
(d) Topo 1



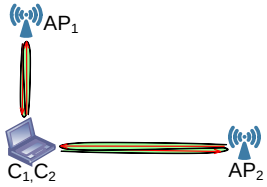
(e) Topo 2



(f) Topo 3

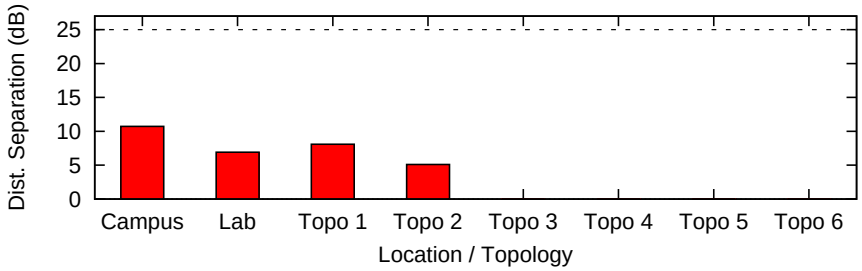


(g) Topo 4

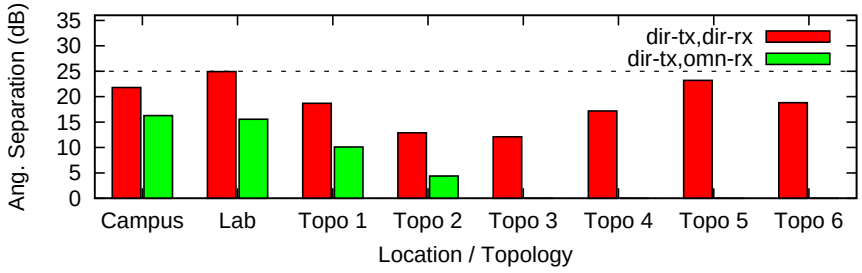


(h) Topo 5

Figure 5.8: Experimental map for evaluating Speed



(a) Distance Separation



(b) Angular Separation

Figure 5.9: Distance and angular separation

two scenarios Campus 3 and Lab 3 are used in Chapter 5.1 to determine the antenna configuration on the clients. Also note that Campus 2 and 3 have the same map, but with different setup. In this experiment, we measure the signal strength from all APs to all clients, with all orientation combinations on the AP and the client. Except for our algorithm, which uses its own MAC protocol, all other approaches use an optimal scheduler that does an exhaustive search on all possible schedules. For each schedule, we use the SINR model to compute the link throughput. The network capacity is the sum of all link throughputs. Since our algorithm involves randomness in terms of which transmission sends the *request-to-send* frame first, we enumerate all possible *request-to-send* sequences and the network capacity is averaged over all sequences.

Separation

Intuitively, the performance of Speed is primarily determined by two factors: the network topology (the location of the APs and clients) and the capability of directional antennas and their orientations. We call these two factors the distance separation and the angular separation. The distance separation is due to the difference between the distance (or more accurately, pathloss) from the client to its own AP and to the interfering AP. The angular separation is due to the ability of a directional antenna to focus its energy on a particular direction.

In Chapter 7, we introduce the notion of separation metric that captures both factors of network topology and directional antenna capability. Generally speaking, a higher separation metric identifies a network with a higher potential for concurrency. In this section, we only use one variant of the separation metric (the link pair metric), which is essentially the sum of SINR values at the receivers. We use the separation metric to understand the performance of various antenna

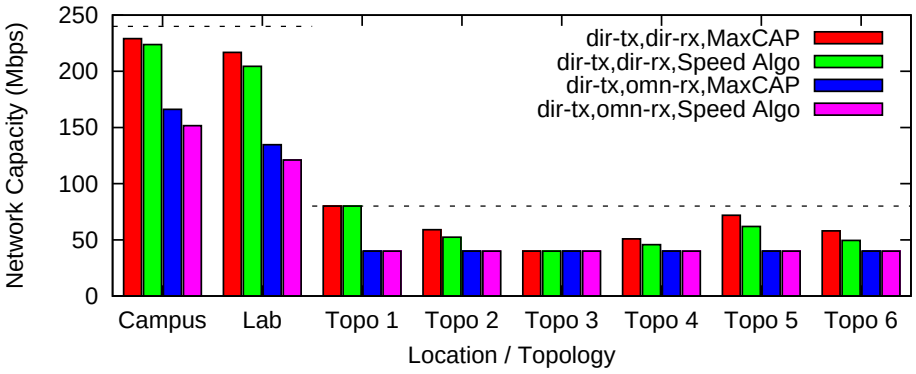


Figure 5.10: Evaluating Speed's algorithm

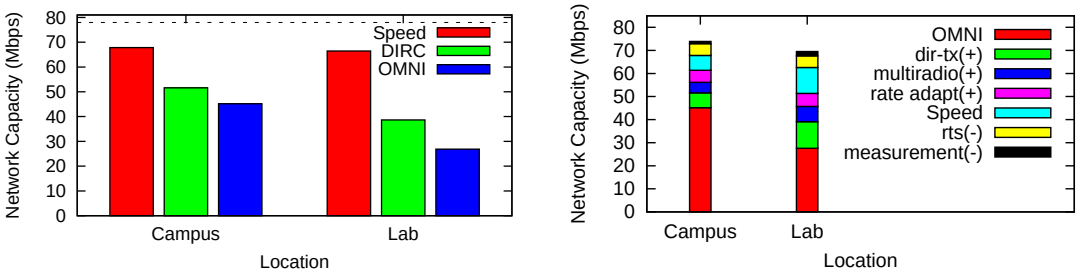
configurations and orientation algorithms.

Figure 5.9 shows both the distance and the angular separation for each topology. Note that the separation shown in these two figures corresponds to the antenna configurations of *dir-tx,dir-rx*, *dir-tx,omn-rx*, and *omn-tx,omn-rx* with the MaxCAP antenna orientation algorithm. The corresponding network capacity is shown across Figure 2.3, Figure 5.5, and Figure 5.10. The dotted line at 25 dB shows the separation needed for concurrent 54 Mbps transmissions. Note that the separation is the sum of both angular and distance separation, and in fact, the distance separation can be defined the same with omni-directional antennas on both APs and clients (where angular separation is 0). Thus, we first obtain the distance separation through the setup with omni-directional APs and clients. Then, by measuring the separation for antenna configurations of *dir-tx,dir-rx* and *dir-tx,omn-rx*, we can identify the angular separation by subtracting the distance separation from the measured separation.

Antenna Configuration

In Chapter 2, we presented the measurement results for four different antenna configuration. Using the separation metric, we can explain why Speed performs close to optimal in the campus and lab scenarios. From Figure 5.9, we observe that the separation is higher than the SINR threshold to decode the 54 Mbps frames. This is achieved by a 10 dB angular separation from clients, about a 10 dB angular separation from APs, and about a 10 dB distance separation.

Next we use the separation metric to explain why *dir-tx&omn-rx* performs poorly when clients are clustered or closely located. First, when clients are clustered, the distance separation is small (the distance separation is 0 for co-located clients). In addition, the angular separation is similarly small when the clients are clustered (angular separation is also 0 for co-located clients if clients are omni-directional). The results show that when the clients are located close to each other, both distance and angular separation will be very small without directionality on clients. Such settings are especially common in settings such as meeting rooms.



(a) UDP (b) UDP Breakdown

Figure 5.11: End-to-End protocol performance: UDP

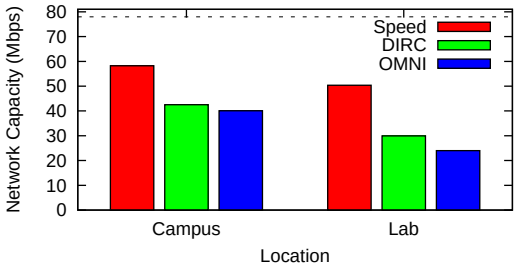


Figure 5.12: End-to-End protocol performance: TCP

Speed's Algorithm

In Section 5.2, we presented the measurement results for two orientation algorithms (MaxCAP and MaxSNR). Next, we present the evaluation of Speed's antenna orientation algorithm. Figure 5.10 shows the performance of Speed's Algorithm, compared with MaxCAP and MaxSNR. The results show that for *dir-tx&dir-rx*, Speed's performance is very close to MaxCAP, i.e., 97% in the campus scenario and 94% in the lab scenario. Speed performs better than MaxSNR. The main reason is that the APs now choose antenna orientations considering the antenna orientations on other transmissions. Thus, the separation of Speed's algorithm will be higher than that of MaxSNR, achieving similar performance to MaxCAP. Note that this is especially true for *dir-tx&dir-rx*, because the separation is high enough and thus the difference in MaxCAP and Speed's algorithm is less apparent.

5.4.2 Implementation

In the implementation of Speed, each directional client is equipped with four 35° patch antennas and an omni-directional antenna. Note that in implementation, unlike in measurement studies, we cannot normalize the signal strength according to antenna gains. And these antennas can usually provide up to 5dB directional gain. Each antenna connects to a separate Atheros 5413 wireless card. Note that Speed assumes a single radio setup on both APs and clients, and we use a multi-radio setup in our prototype to emulate such a system using existing hardware (which only allows for two antenna connectors per card, instead of our target five). Note that this workaround

is not fundamental, i.e., future wireless cards can have multiple antenna connectors, and a single radio is enough. Of course, such a setup is expected to improve network capacity over a single radio setup, but we will factor this effect out later by counting the number of data frames from antennas that are not specified in the *request-to-send* frames. All wireless cards are connected to the PCI slots of a desktop machine through PCIe-to-PCI bridges, and they are all set to the same MAC and IP addresses. We implemented the AP using a Phocus phased array antenna which has 16 directions of 45° beams, plus an omni-directional pattern. Phocus array antennas can switch their orientation in $100\mu s$. All APs and clients operate in ad hoc mode with synchronized clocks.

In order to implement the MAC protocol, we had to work around several limitations of the driver. For example, in Speed's MAC protocol, ideally a node will remove the *request-to-send* frame from its transmit queue when it receives a *request-to-send* from another AP. Unfortunately, the driver does not support this, so we instead allow the frames to be transmitted. In order to ensure ordering, each AP inserts two new fields in the *request-to-send* frame: an order number and a random number. Ties are broken either by a lower order number or a lower random number and the AP resends a new *request-to-send* if it loses the tie. This work-around implementation does not work well with a large number of transmissions, but the limitation is not fundamental; we hope that future driver/firmware implementations provide the ability to drop packets from the transmit queue.

For the timeslot based MAC protocol to work, it is necessary that all frames have the same airtime as a 1500 byte frame transmitted at 54 Mbps frame. This approach incurs padding overhead which is dependent on the payload size and data rate. In the worst case, a frame will incur 1499 bytes of padding overhead at 54 Mbps. Fragmenting a full size 36 Mbps frame into two and a full size 24 Mbps frame into three reduces throughput by 25%. Though this limitation is not fundamental and can be reduced by aggregating frames and delaying transmission, we leave the implementation as future work. In our experiments, we simply work around this problem. In UDP test, we simply truncate the data frames such that the duration of each frame is roughly 222μs. For the TCP test, however, truncating the TCP packet will cause byte loss and retransmission; we work around this issue by using the same data rate on all TCP transmissions.

As mentioned in the previous section, channel dynamics may become more critical due to the strong directionality. The primary ways to detect dynamicity is either at APs, or at the clients. We found detecting throughput drop at the APs much more useful than detecting SINR changes at the clients, and this could deal with all dynamic events we introduced in our testbeds. Thus in our implementation, we only implemented detecting throughput drop at APs. A threshold is needed to trigger measurement updates, which is a tradeoff between responsiveness to dynamicity and measurement overhead. We pick a conservative threshold of 50% drop. We choose a conservative threshold because the threshold is good enough to handle all kinds of dynamic events, i.e., people walking around, blocking the current transmissions, and node mobility in our testbeds. We leave a comprehensive study of how dynamicity affects Speed as future work.

5.4.3 System Level Evaluation

The goal of this experiment is to evaluate the implementation and the end-to-end performance of the system.

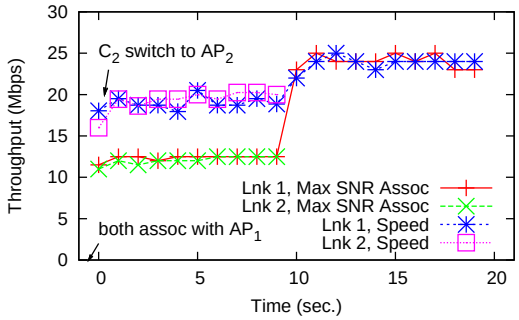
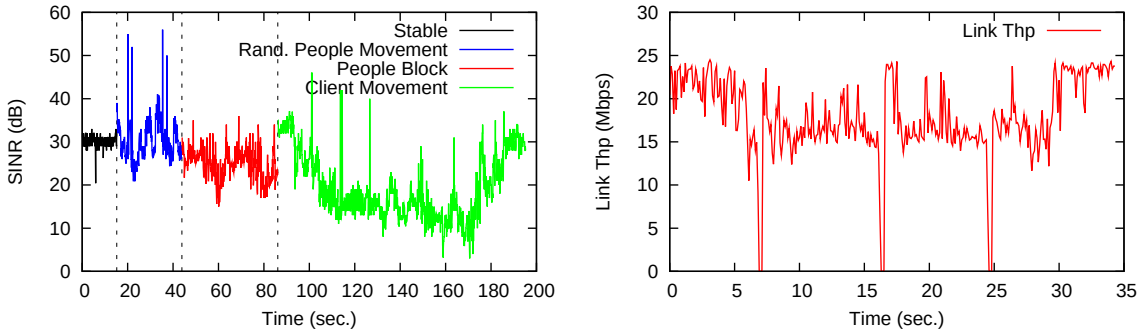


Figure 5.13: Evaluating association in Speed

UDP & TCP Performance

We evaluate both the UDP and TCP performance of Speed in the campus and lab scenarios (Figures 5.8(a) and (b)). We placed six clients C_1 to C_6 as indicated on the map, and placed three directional APs in various locations. Note that in the measurement-based study, we are able to emulate six APs by putting one AP at six locations at different times, but in this evaluation, we are limited to the three phased-array antennas we have. In the campus scenario, the APs are in $(1, 2, 3), (1, 3, 5), (2, 3, 5)$, respectively; and in the lab scenario, the APs are in $(1, 4, 6), (1, 2, 5), (1, 2, 3), (2, 3, 5), (1, 4, 5)$, respectively. For each AP location, we activate all possible client combinations, and present the mean capacity from all combinations. For each setup, the experiment runs for 1 minute, and the results are averaged over 3 runs. Since we did the experiments in the evening, and all 3 runs are experimented back-to-back, the results from 3 runs are very similar. Thus we do not show confidence intervals in the graph. Figure 5.11 and Figure 5.12 show the UDP and TCP performance for Speed, DIRC (as described in Chapter 4), and OMNI (omni-directional APs and clients). The results show that in the lab scenario, Speed improves UDP performance over DIRC by 100% and over OMNI by 127%, while Speed improves TCP performance over DIRC by 56% and over OMNI by 93%. In the campus scenario, Speed improves UDP performance over DIRC by 31% and over OMNI by 50%, while TCP performance is improved by 36% over DIRC and by 45% over OMNI. The reason that the improvement in the lab scenario is much higher than that in the campus scenario is that the distance separation in the campus scenario is higher, thus the performance of DIRC and OMNI in the campus scenario is much better. TCP performance of Speed degrades more than expected, especially in the Lab 2 scenario, because we fixed the transmit rate for the TCP experiments.

Figure 5.11(b) shows the breakdown of gains and overheads for UDP performance for the Lab 2 and Campus 2 scenarios. In the Lab 2 scenario, directional antennas at the transmitter contribute 42% improvement in network capacity over omni-directional antennas. Adding directional antennas on clients increases the performance by 85% over the directional transmitter scenario. Enabling frames to be received across all the radios in Speed's setup improves performance by 9%. The impact of rate adaptation is also important where we see that disabling rate adaptation will lead to a 11% reduction in performance. The figure also shows the protocol performance overhead in the same stacked bar. The overhead of *request-to-send* control frames is 8% of the airtime and measurement updates contribute only 3% of the total airtime.



(a) Impact of dynamic events

(b) Handling Mobility

Figure 5.14: Evaluating dynamics in Speed

Association

To evaluate the association process, we use two directional transmissions, constructed as in Topology 5. We emulate the following 20-second scenario with the two transmissions: 1) in the first 10 seconds, both clients are active, and 2) in the next 10 seconds, C_2 leaves the network. APs will send UDP traffic to all associated clients. Figure 5.13 shows this process for Speed and for MaxSNR association. In the first 10 seconds, both clients associate with the same AP AP_1 for MaxSNR association. While in Speed, C_1 associates with AP_1 and C_2 associates with AP_2. Due to this association, the link throughputs from Speed are much improved over that from MaxSNR association. Also note that even though C_2 associates with an AP that does not have the strongest signal, the link throughputs of the two links are comparable. This is because of the randomness in the Speed MAC protocol, i.e., for each timeslot, both links can reserve the channel before the other. After C_2 leaves the network, the link throughputs from both mechanisms are comparable.

Dynamics

In this experiment, we first evaluate the effect of channel dynamics on Speed's performance in the case where Speed does not try to adapt. We enable two transmissions in the research lab, $AP_1 \rightarrow C_1$ and $AP_3 \rightarrow C_3$, and measure the SINR level at C_1. (Here we show SINR values instead of throughputs to portray the real time evolution in client performance.) We measure the following: 1) a baseline SINR where the environment is reasonably stable (i.e., no people movement or mobility); 2) SINR where a person circles around the room; 3) SINR where a person moves back and forth to block the LOS path between AP_1 and C_1; and 4) SINR when we move C_1 to the location of C_2 then to C_3 and finally back to C_1. Figure 5.14(a) shows the SINR levels on C_1 during this process; the dotted vertical lines indicate the transition from one experiment to the next. The results demonstrate the impact of dynamicity on the SINR level, below and even above the baseline. The SINR level rarely drops so low that the link cannot sustain some (lower) data rate. Note that though we did not show the SINR levels at the second receiver C_3, they remain relatively stable during this process.

Next, based on dynamicity case (4) above, we show how Speed can handle node mobility,

which is the most dynamic scenario. Recall that in this case C_1 is moved from the location of $C_1 \rightarrow C_2 \rightarrow C_3 \rightarrow C_1$, while both transmissions $AP_1 \rightarrow C_1$ and $AP_3 \rightarrow C_3$ are active. Figure 5.14(b) shows the link throughput on $AP_1 \rightarrow C_1$ over time, and during this process, three conflict graph updates are triggered at $t = 6.9s, 16.3s, 24.6s$.

5.5 Summary and Discussion

In this chapter, we presented the design, implementation, and evaluation of Speed. Speed is a directional antenna system designed for future wireless networks with both directional APs and clients. Similar to DIRC, Speed also uses the SINR model and the conflict graph to determine antenna orientations. Speed is different from DIRC (Chapter 4) in that the clients are also directional, the clients consider which AP to associate with, and that the MAC protocol is distributed. These differences cause several challenges in designing the Speed system.

We first presented a detailed exploration of indoor directional antenna systems that maximize spatial reuse along two dimensions: antenna configuration and control. Based on this exploration, we designed and implemented Speed. The antenna control system uses conflict graphs and traffic load estimates to overcome the unique challenges of antenna orientation, MAC protocol, and client-AP association posed by a network of directional APs and clients. Our evaluation of Speed in two indoor testbeds show that Speed can indeed maximize spatial reuse by increasing network capacity by 31% and 100% over existing solutions.

Speed also has several limitations. First, adapting the timeslot based MAC protocol to small TCP transfers, e.g., web browsing, is challenging. We will discuss this topic more in the next chapter, and several mechanisms may be used to improve the performance under the light and bursty network usage, including reducing timeslot size, allowing APs to send frames to multiple clients in a single timeslot, allowing APs to explicitly cancel a reservation if it is not currently utilized, etc. Second, while four patch antennas can be deployed on certain devices such as laptops, smaller devices like smartphones pose additional challenges in terms of antenna deployment. Also, Speed APs are equipped with phased-array antennas, but in applications such as data transfer between a camera and a laptop, phased-array antennas do not fit on either. We leave the exploration of smart antennas that can leverage spatial reuse on these devices as future work.

Chapter 6

Opera: Spatial Reuse for Omni-directional Antenna Networks Through Power Control

In Chapter 4 and Chapter 5, we have presented two directional antenna systems that optimize indoor wireless spatial reuse. These two systems require significant deployment efforts. Also at the same time, there is always a chance that directional antennas never really take off, for example, due to high cost (existing commercial directional antennas, such as Xirrus sectored WiFi array [8], Phocus array [3], or Ruckus BeamFlex [5], costs several thousand dollars). Thus, it is also important to apply the notion of spatial reuse to existing omni-directional antenna networks.

We address this issue by proposing Opera [73], a power control system that is designed to optimize spatial reuse for omni-directional antenna networks. In fact, many projects [11, 32, 42, 56, 60, 64, 81, 82, 83, 84, 85, 86, 92, 95, 104, 106, 107, 128] have explored adjusting transmit power and/or CCA threshold to improve network capacity and/or reduce energy consumption in wireless networks. Note that the goals of improving capacity and reducing energy consumption are not always compatible. Our focus in Opera is to optimize network capacity by improving wireless spatial reuse. In Chapter 2.4.3, we showed that existing solutions that either make uncoordinated decisions or rely on CSMA-based MAC protocols, only work well in homogenous networks with predictable interference. They fail to maximize the spatial reuse in chaotically deployed networks [11, 40]. In Chapter 1, we mentioned that the unique characteristics of the application scenario is that the receive patterns from the sender is unpredictable. As a result, interference varies widely even among wireless nodes that are in the same region. Figure 6.1 shows an example setting where interference within a single cell significantly differs depending on the location of the clients. In this setting, C_3 experiences stronger interference than C_1 and C_2. We show other example later with varying interference levels for nodes across cells. Based on measurements in five different locations, we show that similar scenarios occur in practice.

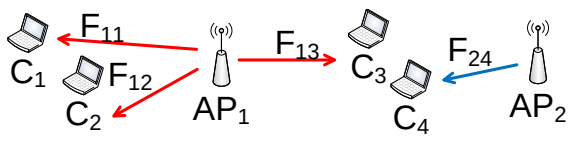


Figure 6.1: Chaotic deployment leads to variable node densities and unpredictable interference

Protocol	TxPower	MAC	Granularity	Interf. Model
[11, 95],etc.	MinPC	default CSMA	per-link	range-based
[64]	OptPC w/ well placed APs	CCA tuning	per-link txp, same CCA	SINR
[81]	OptPC	CCA tuning	per-cell	SINR
Opera	OptPC	conflict graph based	per-link	SINR

Table 6.1: Summary of power control systems in various dimensions

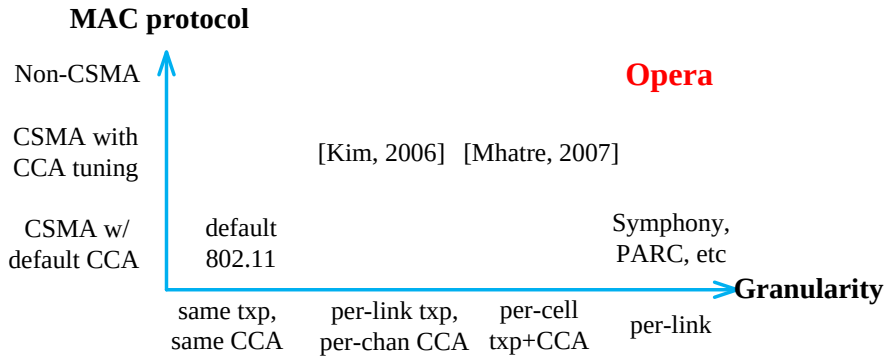


Figure 6.2: Categorization of existing distributed power control solutions

In this chapter, we propose Opera, a practical, distributed and fine-grained power control protocol that optimizes spatial reuse, even in chaotic deployments. Opera system includes an algorithm that tunes power level on a per-link basis and a channel reservation based MAC protocol that works well with the power control algorithm. We make contributions in two areas. First, existing designs for distributed power control have used the CSMA MAC protocol, either with the default CCA threshold, or CCA tuning. We show that CSMA MAC protocols interact poorly with transmit power control and, thus, cause these designs to perform poorly. In Opera, we propose a system that uses a MAC protocol that is based on channel reservations (similar to Speed's MAC protocol presented in Chapter 5). This MAC is able to identify the active senders of the network and coordinate their choices of power levels. Second, since interference in chaotic networks is unpredictable, no assumptions can be made on the interference level at each client. Our system rapidly and accurately characterizes the environment that the system operates in and automatically calibrates the SINR thresholds on each wireless card. Our experimental results show that Opera performs well in practice, and can improve over existing solutions by 40%.

6.1 Related Work

Figure 6.2 categorizes existing power control solutions and Opera in a two dimensional design space: the MAC protocol and the granularity of tuning, Our protocol, Opera, is the only system that uses a non-CSMA MAC layer and tunes the power level on a per-link basis. We will show in Chapter 6.2.1 that this point in the design space is capable of achieving higher spatial reuse than

other points in the design space. Note that these two dimensions do not fully define a system and we highlight other key design dimension choices for both Opera and other related work below. Also note that we changed one dimension in this space from the two dimensions presented in Figure 1.2 because the granularity is a better dimension for power control to distinguish Opera from prior efforts.

6.1.1 Dimension 1: MAC Protocol

Most power control solutions use the 802.11 MAC protocol, CSMA, with default CCA threshold. In these solutions, senders reduce the power level to the minimum necessary to decode frames at the receivers [11, 32, 56, 60, 82, 83, 84, 85, 86, 92, 95, 104, 106, 107, 128]. The minimum power is chosen for two reasons. First, it minimizes interference at other nodes. However, it also minimizes the signal strength at the intended receiver, making the links susceptible to external interference, which in most cases is suboptimal. Second, it minimizes energy consumption. In Opera, we do not focus on energy consumption, but we expect Opera's energy consumption will be no more than that of standard 802.11 since it may use lower transmit power levels. The biggest problem with the CSMA MAC protocol with default CCA thresholds in these approaches is that using power control can cause unfairness or link asymmetry, i.e., links that use low power levels are more likely to be interfered with and less likely to be heard by interferers. This asymmetry problem has been widely observed in previous work [11, 32, 60, 82, 83, 84, 86, 95, 104, 106]. Several proposed designs do deal with this problem. For example, Symphony [95] detects this asymmetry when throughput drops below expected values and responds by triggering a power increase.

There are power control systems that tune CCA thresholds along with txpower levels [64, 81]. In Chapter 6.2.1, we show that CCA tuning also interacts poorly with power control and, thus, these systems cannot maximize spatial reuse in chaotic networks.

6.1.2 Dimension 2: Granularity

Another dimension is the granularity of tuning. For example, [64] assigns the same CCA threshold to every node in the network, and [81] uses the same power/CCA configuration for all the nodes in the same cell. Many other solutions [95] tune system parameters on a per-link basis. In [85], the authors conclude that coarse-grained approaches are asymptotically optimal, but we show in Chapter 6.2.1 that spatial reuse can still be greatly limited by the "worst" client in coarse-grained approaches.

Table 6.1 summarizes different systems in more dimensions.

6.1.3 Cellular Networks

Power control has been deployed in cellular networks to achieve cellular frequency reuse. As mentioned in Chapter 1, the cellular networks operate in licensed frequency bands: the network provider carefully plans the locations of the cellular towers and thus the interference is very predictable. Thus this type of networks are very different from our application scenarios, and their solutions do not apply.

On the other hand, the recent technology of femtocells [28, 29, 30, 57, 124, 125], i.e., small cellular base stations that can be plugged in in residental or small business settings, may change the landscape of cellular networks dramatically because the placements of femtocells are unplanned (similar to AP placement in these settings). Femtocells, however, are still different from wireless networks because they operate in licensed frequency bands and they are connected to the wired network. Note that current research (specifically power control) on femtocells is still in its early stage, and we believe that research in this area will become a bridge between the cellular networks and wireless networks, and femtocell protocols will include ideas from both areas.

6.2 Opportunities and Challenges

In Chapter 2, we have shown the performance benefits of power control in two indoor scenarios (Campus 2 and Lab 2), and that default CSMA does not work well through an example (Figure 2.5). We do a similar evaluation in this section.

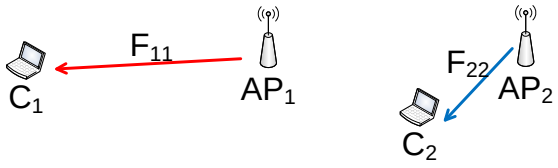
First, we explain why existing different categories of solutions (from the previous section) fail to maximize spatial reuse. Then, we use measurements collected in five different locations to show that Opera does improve network capacity over existing solutions in practice.

6.2.1 Problem with Existing Solutions

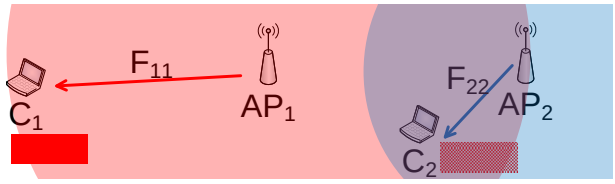
First, we show that any solution that uses a CSMA MAC protocol with the default CCA threshold does not maximize network capacity. As discussed in the previous section, power control can lead to undesirable link asymmetry and existing solutions propose to increase power level when link asymmetry is detected [95]. However, this power increase may trigger other senders to increase their txpower as well. Eventually, these increases may result in all senders hearing each other and eliminate any possible spatial reuse. This is undesirable since the goal of the system is to increase spatial reuse.

Figure 6.3(a)-(c) shows an example with two transmissions, where C_1 is relatively farther away from AP_1 than C_2 from AP_2. Each horizontal bar in the figure shows the link throughput (dark part) vs. the optimal link throughput that can be achieved (dark plus light parts). In Figure 6.3(b), if nodes use the minimum txpower, AP_1 cannot hear and will not defer to AP_2's transmissions and, thus, AP_1's transmissions will interfere with C_2's reception. As shown in Figure 6.3(c), triggering the power level increase at AP_2 solves the problem of link asymmetry However, it prevents any spatial reuse because both AP_1 and AP_2 can hear and defer each other.

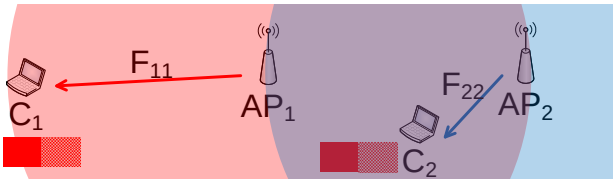
Second, we show that for existing coarse grained power control solutions, spatial reuse can be greatly limited by the "worst" client, and thus are suboptimal. One example of this limitation is shown in Figure 6.3(d)-(f). In the example, receiver C_3 is in a poor location since flow F_{13} interferes with F_{24}, but all other flow transmissions can transmit simultaneously. In this example, concurrent transmissions can only happen without incurring starvation when using a per-link granularity protocol for the following reasons. First, without per-link txpower, as in [81], the same power level will be used by AP_1 to all its clients C_1, C_2, C_3 (i.e., for flows F_{11}, F_{12}, F_{13}). And AP_2 uses the same CCA threshold at all times and is unable to treat any of these three flows differently: AP_2 can either use a high CCA that causes C_3 to starve (Figure 6.3(e)), or it can use



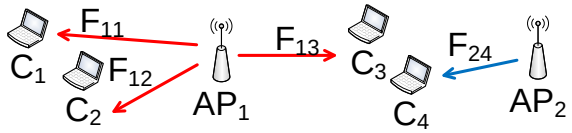
(a) Example 1



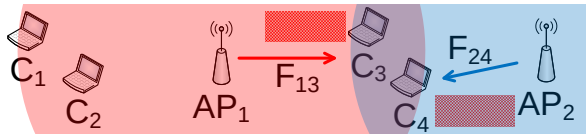
(b) AP_2 uses low power: causes link asymmetry



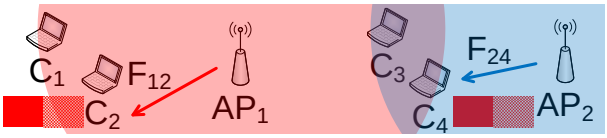
(c) AP_2 uses high power: no spatial reuse



(d) Example 2



(e) $S1$ and $S2$ do not defer: cause collision



(f) $S1$ and $S2$ defer: no spatial reuse

Figure 6.3: Two motivating example

a low CCA that wastes the concurrent transmission opportunities with F_{11}, F_{12} (Figure 6.3(f)). Second, without per-link CCA or other equivalent MAC protocols, as in [64], AP_1 can either use a high CCA that let F_{13} and F_{24} interfere with each other, causing F_{13} to starve (Figure 6.3(e)), or it can use a low CCA which again wastes the spatial reuse opportunities (Figure 6.3(f)). This example illustrates that the performance of coarse-grained protocols is limited by the "worst" client.

Note that in [64], the authors conclude that tuning txpower or CCA threshold has similar effects, which implies that CSMA with the same CCA threshold will work just fine if txpower levels are tuned. However, their conclusion only applies to the scenario presented in their analysis: a network where all APs are uniformly distributed (honey-grids). In contrast, we are interested in chaotically deployed networks.

Another observation made from the two examples in Figure 6.3 is that the key requirement for the MAC protocol is to identify the active senders. For example, the reason that the per-cell solution fails to maximize network capacity in Figure 6.3 is because sender AP_2 cannot identify which receiver AP_1 is talking to. Even with per-link CCA tuning, the senders can only estimate the active senders by measuring energy level instead of identifying them. This is also why CSMA-based approaches do not work well.

6.2.2 How About in Practice?

The two examples in Figure 6.3 illustrate why existing power control solutions fail to maximize network capacity in chaotic networks. Here we evaluate these existing solutions in five locations, based on collected measurements, and show that examples do show up in practice. The experimental maps for two locations are shown in Figure 5.8(a)&(c). (the maps for the other locations are unavailable), and the details of the experimental setup are described in Section 6.5.

We compare several solutions, including:

1. default 802.11,
2. Symphony [95] that tunes power levels and uses CSMA with default CCA threshold,
3. CENTAUR [109] that uses a centralized/smarter MAC protocol but does not tune power levels,
4. per-cell tuning of txpower and CCA thresholds [81],
5. per-link tuning of txpower and CCA thresholds,
6. Opera (per-link txpower tuning with a non-CSMA MAC protocol), and
7. Optimal configurations obtained using a optimal scheduler through exhaustive search.

Figure 6.4 shows the network capacity from each solution in each location. The results show that Opera outperforms existing solutions in all five locations, suggesting that the examples do show up in practice. Also, Opera performs reasonably close to the optimal exhaustive search solution. We present a detailed explanation of these results in Section 6.5.1. Notice the per-link tuning of txpower and CCA thresholds approach. The performance of this approach is worse than, but close to, that of Opera. However, we will talk about why this approach is not desirable in Chapter 6.3.2.

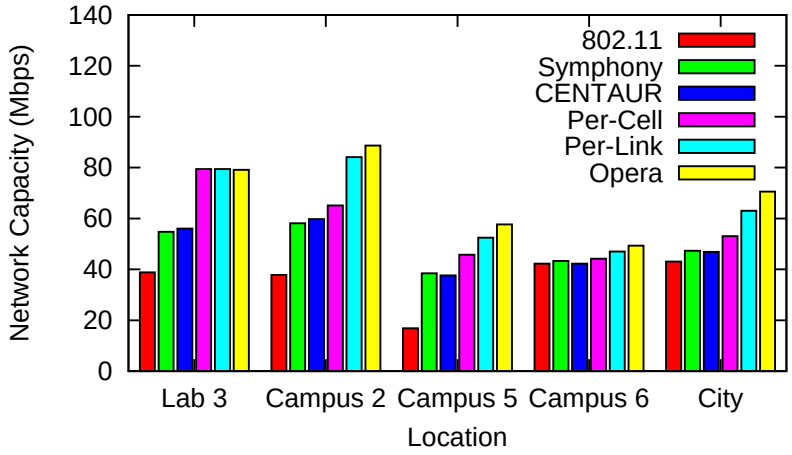


Figure 6.4: Network capacity of various algorithms

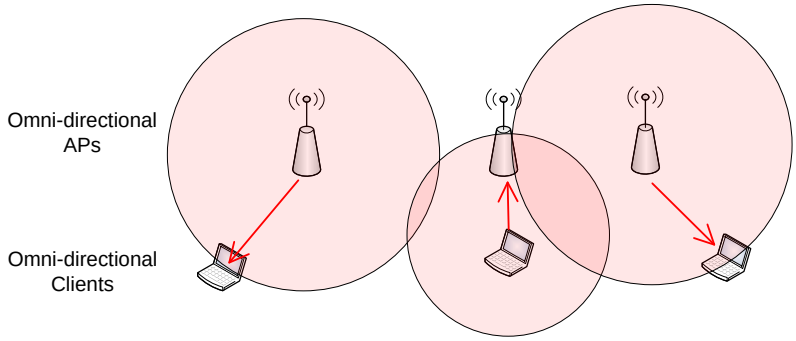


Figure 6.5: Opera system overview

6.3 Opera Protocol

The goal of the Opera design is to optimize spatial reuse in chaotic networks, and the approach we take is per-link txpower tuning and non-CSMA MAC protocol. The challenges of designing this protocol are threefold. First, finding the optimal power levels is conditional NP-hard (see Chapter 6.6 for proof), and the problem is further complicated by the fact that Opera is a distributed protocol. We propose to use a heuristic algorithm that works well in practice. Second, question remains what distributed MAC protocol can satisfy the requirement of identifying active senders. We evaluate two candidates and pick one that works better with our power control algorithm. The final challenge is how to make the protocol practical, especially how to collect accurate measurements. In this section, we describe the Opera protocol and how it addresses these three challenges: the MAC protocol design, the power level tuning design and the measurement collection protocol. Figure 6.5 shows Opera's system overview.

6.3.1 Overview

In this section, we present an overview of the Opera protocol. First, in order for Opera to work in densely and chaotically deployed wireless networks, with highly variable node densities and unpredictable interference, Opera relies on measurements to characterize the environment it operates in. Opera estimates the impact of interference using the SINR model with multiple data rates, as described in Chapter 3. Note that in the Opera design, we assume a single SINR threshold to simplify presentation, but in the Opera implementation (Chapter 6.4), we assume a curve of SINR vs. loss rates instead of a threshold, and calibrate this curve on each card.

Second, as illustrated from the examples in Figure 6.3, the key requirement for the MAC protocol is to identify the current transmissions. Also, in Chapter 2, we saw that CSMA based approach is not a suitable choice. In Chapter 6.3.2, we present three possible design choices and choose the one that works better with Opera's power control algorithm.

Third, choosing the optimal power levels on a collection of active senders is conditional NP-hard (see Chapter 6.6 for proof). In Chapter 6.3.4, we propose a heuristic power control algorithm that works well in practice. In Chapter 6.5.1, we provide some insights into why it works well and how it works in different scenarios.

6.3.2 MAC Protocol

As mentioned previously, the key requirement for the MAC protocol is to identify current transmissions.

We first examine why CCA tuning is not a desirable choice. The biggest hurdle is that tuning CCA thresholds along with txpower levels is very difficult, both inherently and practically. The first problem is that the CCA thresholds need to be considered together with txpower levels in the problem formulation that maximizes network capacity. For example, changing txpower levels cause changes in received energy levels and, thus, changes the optimal CCA thresholds. The change of CCA thresholds causes on/off behavior of interfering links and, thus, affects the optimal txpower levels. These interactions make simultaneously determining the optimal txpower levels and CCA thresholds on a per-link basis and in a distributed fashion both difficult and/or suboptimal. As a result, these systems can suffer from undesirable link asymmetry and thus node starvation or, by keeping the product of power level and CCA threshold a constant, they can prevent link asymmetry [81] but suffer from limited spatial reuse. In fact, we did design and implement a system that depends on per-link txpower control and CCA tuning [74]. However, due to the reasons mentioned above, we decided to switch the MAC protocol in the Opera system.

The second problem with CCA threshold tuning is that it is difficult to implement. Most existing wireless cards do not support CCA threshold tuning at all. Others cards tie CCA thresholds to receiver thresholds and on these cards raising CCA thresholds raises receiver thresholds. This prevents the card from receiving packets from nodes other than their intended receivers, causing deafness problem. Such behavior could be especially undesirable for APs when a new client joins the network. Also as mentioned earlier, existing hardware does not support per-packet CCA tuning, and does not work very well with non-protocol-compliant nodes. Due to this reason, there is no implemented system for per-link power control and CCA tuning, and [81] only supports operations with offline measurements. Last, the performance of this approach can be

worse than Opera, e.g., up to 35%.

Then let us examine two pieces of existing work on non-CSMA based MAC protocols: CMAP [121] and Speed as described in Chapter 5.

In CMAP, a sender can identify transmissions by overhearing their frame headers, and based on this information, it can decide to transmit if it does not interfere with overheard transmissions. CMAP is an asynchronous approach where each sender can initiate transmissions at any time. As described in Chapter 5, Speed's MAC protocol is based on timeslots and timeslot reservations. In Speed, a sender can reserve the next timeslot by sending a reservation message and all senders can identify the active senders for the next timeslot by overhearing these messages. A sender can reserve a timeslot if it does not interfere with existing reservations for that timeslot. Speed's MAC protocol is a synchronous protocol, i.e., all senders that have reserved the timeslot transmit simultaneously.

The performance of Speed's MAC protocol is better than that of CMAP's due to its synchronicity. The following example illustrates the performance benefits. Suppose there are two non-interfering flows running CMAP, F_1 and F_2, and F_1 starts transmitting first. Then, F_2 hears about F_1, and since they do not interfere, F_2 can start transmitting immediately. When F_1 finishes its transmission, F_1 can tell there is another transmission but cannot identify the source or destination. Thus, F_1 has to wait until the end of F_2's transmission, wasting spatial reuse opportunities. If both flows overlap randomly in time by 0-100%, CMAP wastes an expected number of 50% of the spatial reuse opportunities. This problem does not occur in a synchronous protocol, where all senders reserve timeslots and transmit synchronously. For example, in the previous example, with a synchronous protocol, both F_1 and F_2 will start and end their transmissions at the same time. Thus, in this Opera, we use the synchronous protocol.

Speed's MAC protocol is based on timeslots and timeslot reservations. The most desirable property of this MAC protocol is that it allows multiple distributed senders to transmit selected data traffic concurrently. Speed uses CSMA to prevent collisions for other traffic, such as timeslot reservations, TCP ACKs from clients, traffic from external APs and clients. Note that Speed's MAC protocol is designed for directional antenna networks. In Opera's variant of the Speed design, in each timeslot, instead of picking antenna orientations, each AP runs the power control algorithm and reserves the timeslot if it does not interfere with existing reservations. We also change some parameters of the Speed MAC protocol to meet the requirement of Opera. For example, Opera uses a 10ms timeslot size (instead of 20ms) and updates measurements every 1 second (instead of 5s).

6.3.3 A Two-Flow Example

To gain insight into choosing the right power levels, we begin by considering a simple scenario (Figure 6.6) where S_1 transmits to R_1 and S_2 transmits to R_2 (refer to Chapter 2.1 and Chapter 2.2.3 for notations). Note that we do not assume pathloss obeys the triangle inequality in our algorithms or protocols. The SINR at receivers R_1 and R_2 are $SINR_1 = P_1 - L_{11} - P_2 + L_{21}$ and $SINR_2 = P_2 - L_{22} - P_1 + L_{12}$, respectively, where P_i is the transmit power level from S_i to R_i, and L_{ij} is the path loss from S_i to R_j $(i, j \in \{1, 2\})$. Note that independent of the transmit power levels, we have $SINR_1 + SINR_2 = L_{12} + L_{21} - L_{11} - L_{22}$. Power control essentially allocates this sum between the two transmissions, i.e. increasing $SINR_1$ will decrease

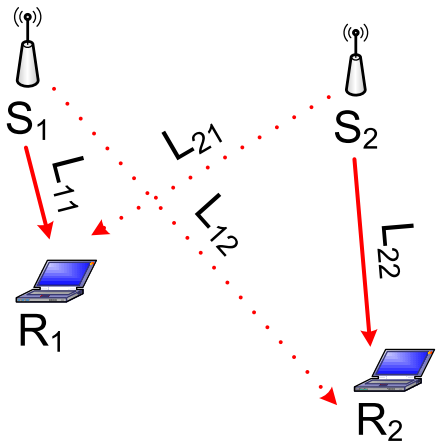


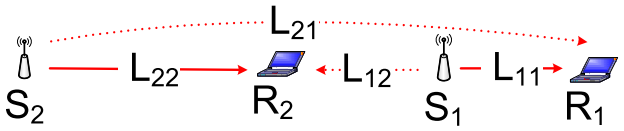
Figure 6.6: A general two omni-directional flow topology

$SINR_2$. In order to enable concurrent transmission, we need both $SINR_1 \geq SINR_{thrsh}$ and $SINR_2 \geq SINR_{thrsh}$. Note that it may not be possible to satisfy both constraints, so concurrent transmission may be impossible.

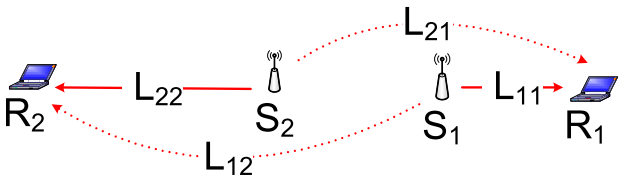
We now consider what happens when all nodes use the same power (i.e., NoPC, as devices all use the default txpower) or when they use the minimum power level to reach the receiver (i.e., MinPC, as proposed in many solutions). With NoPC, we have $SINR_1 = L_{21} - L_{11}$, and $SINR_2 = L_{12} - L_{22}$. Thus, in four typical topologies shown in Figure 6.7, NoPC performs well in scenario (b) & (c) because $SINR_1 \approx SINR_2$ but poorly in (a) & (d) because $SINR_1 >> SINR_2$. With MinPC, we have $SINR_1 = L_{21} - L_{22}$, and $SINR_1 = L_{12} - L_{11}$. Thus, MinPC performs well in scenario (a) & (d) but poorly in (b) & (c). Intuitively, if the sender is far away from the receiver, the transmission is more likely to be penalized using NoPC because its received signal strength is relatively weak. Using MinPC has the opposite problem. If the sender is close to its receiver, it is more likely to be penalized because the interference level is likely to be high.

In order to illustrate this, we constructed the scenarios shown in Figure 6.7 on a 4-node testbed and used it to compare the performance of NoPC, MinPC, and OptPC algorithms. The description of the algorithms can be found in Chapter 2. We first obtained a baseline throughput by having only one of the sources transmitting. We then ran experiments with both sources active, using the three transmit power settings described above. For this two flow scenario, we also emulate the OptMAC behavior by measuring the performance both with the default 802.11 CCA threshold and a high CCA threshold that prevents transmitting nodes from deferring to each other.

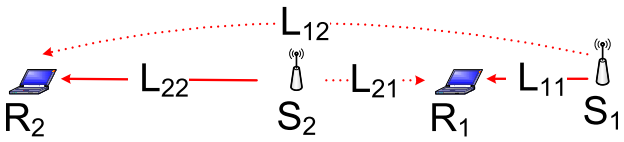
The results, as a percentage of baseline, are shown in Table 6.2. We see that OptPC can enable concurrent transmissions in both scenarios, while neither NoPC nor MinPC work well in both scenarios. Specifically, OptPC alleviates the hidden terminal problem caused by MinPC in scenario (c), and the hidden terminal problem caused by NoPC in scenario (a).



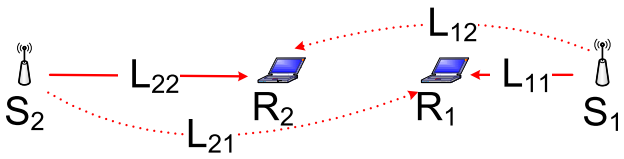
(a) Topo A



(b) Topo B



(c) Topo C



(d) Topo D

Figure 6.7: Four typical topologies

Scenario	PC Algorithm	Thruput 1 (%)	Thruput 2 (%)	Capacity
Figure 2.5(c)	NoPC	81.4	96.6	178.0
	MinPC	11.2	102.0	113.2
	OptPC	94.3	90.0	184.3
Figure 2.5(a)	NoPC	20.8	98.8	119.6
	MinPC	98.8	85.9	184.7
	OptPC	94.1	93.3	187.4

Table 6.2: Evaluation of power control algorithms on a 4-node testbed

6.3.4 Selecting Power Level

Unfortunately, choosing the optimal power levels is conditional NP-hard, we propose a heuristic power control algorithm that works well in practice. The huristic algorithm is based on the insight obtained from the two-flow power control example described in Chapter 2. The algorithm (shown in Algorithm 7) is run on every sender for each timeslot. The algorithm assumes that every node has knowledge of the network topology (i.e., path loss from APs to clients). Opera incorporates techniques for collecting these measurements .

The core of our power control algorithm is a heuristic where each sender reserves the channel if it does not interfere with existing reservations. In the algorithm, the sender first estimates the capacity that can be achieved if the sender defers (line 1). Algorithm 6 shows the algorithm that estimates network capacity given the timeslot reservations, based on the SINR model. Then, each AP iterates through each transmit power level (p) and each data rate (rt), the maximum network capacity can be achieved (line 4-14). Note that there is usually a range of transmit power levels ($[txp\_st, txp\_ed]$) on the sender that can achieve maximum capacity (line 11-12). In our algorithm, we pick the middle level of that range (line 16). This choice is based on the insight gained from the two-flow scenario, i.e., the difference between the power levels from every two senders need to be within a range, so that the probability of concurrent transmissions with the subsequent senders is maximized. For example, suppose the range of txpower levels is $[4, 20]$dBm and if the sender uses 4dBm, it leaves no slack to accommodate external interference from a potential concurrent sender; and if the sender uses 20dBm, it may cause high interference towards a subsequent potential transmission and prevent it from reserving the channel.

In Chapter 6.5.1, we show that our algorithm works well in five locations and provide some insights into why.

Algorithm 6: Estimate network capacity $calc\_cap$

Input: schedule $R = \{(tx, rx, pow, dr)\}$
Result: network capacity cap

1 $cap \leftarrow 0$
2 **foreach** $r = (tx, rx, pow, dr) \in R$ **do**
3 $\quad sig \leftarrow S(tx, rx) + pow$
4 $\quad max\_intf \leftarrow -95$
5 $\quad$ **foreach** $r' = (tx', rx', pow', dr') \neq r \in R$ **do**
6 $\quad\quad intf \leftarrow S(tx', rx) + pow'$
7 $\quad\quad$ **if** $intf > max\_intf$ **then** $max\_intf \leftarrow intf$
8 $\quad$ **end**
9 $\quad sinr \leftarrow sig - max\_intf$
10 $\quad cap \leftarrow cap + sinr\_to\_thruput(sinr, dr)$
11 **end**

Algorithm 7: Power control algorithm

Input: Existing schedule R, current destination rx,
Result: If transmit, txpower txp and data rate dr; If defer, nothing

$cap\_defer \leftarrow calc\_cap(R)$
$cap\_simul \leftarrow 0, txp\_st \leftarrow 0, txp\_ed \leftarrow 0$
$cap \leftarrow \{0\}, drate \leftarrow \{0\}$
for $p = 1 \rightarrow MAX\_TXPOW$ **do**
    for $rt \in DATA\_RATES$ **do**
        $cap\_curr \leftarrow calc\_cap(R \cup (self, rx, p, rt))$
        if $cap\_curr > cap[p]$ **then**
            $cap[p] \leftarrow cap\_curr$
            $drate[p] \leftarrow rt$
        end
        if $cap[p] = cap\_simul$ **then** $txp\_ed \leftarrow p$
        else if $cap[p] > cap\_simul$ **then** $txp\_st \leftarrow p; cap\_simul \leftarrow cap[p]$
    end
end
if $cap\_simul > cap\_defer$ **then**
    $txp \leftarrow (txp\_st + txp\_ed)/2$
    $dr \leftarrow drate[txp]$
end

Data Rate and Adaptation

Note that our algorithm also considers transmit data rate choices in the Algorithm 7 (line 5). Although the data rate is picked for each sender in each timeslot, it can be adapted to handle external interference and to accommodate protocol failure in certain scenarios. Opera can work with multiple rate adaptation algorithms, but in our implementation we use the simple auto rate fallback mechanism.

Measurement Collection

In order to run the power control algorithm described in Algorithm 7, wireless nodes need information about received signal strength from all APs to all clients, i.e., $S(tx, rx)$ in Algorithm 6. Thus, each client records the received signal strength from all APs. Such measurements come from two sources: normal traffic and measurement traffic. First, any client that receives a frame from any AP records its received signal strength; this approach incurs negligible overhead. Although this approach works well for the received signal strength from an AP to its intended receiver, it has two limitations for other measurements: 1) an AP that uses a low transmit power level reduces the probability of reception at unintended clients, and 2) if multiple senders are active in one timeslot, certain measurements cannot be updated (a client that is actively receiving cannot record signal strength from other APs). Thus, measurement traffic is also necessary. Opera has all APs periodically transmit a number of frames one at a time, with highest possible txpower level and the lowest data rate. Clients record the RSSI readings from every AP and send this information back to all the APs. In Opera, measurements are updated once every one second. Thus for 4 APs, it takes 20 ms to take the measurements (5 ms for each AP), and this incurs 2% (20 ms / 1 s) overhead.

6.4 Operation and Implementation

In this section, we present Opera implementation and how it operates. The main issue we deal with in implementation is automatic calibration, which is a critical to obtaining the accurate SINR measurements that Opera relies upon.

6.4.1 Automatic Calibration

In previous section, we presented the power control algorithm and the measurement collection process. The core of the algorithm, i.e., estimating link throughputs using the SINR model, relies on accurate SINR thresholds on receivers. While SINR thresholds are different for different data rates, we only present the calibration of the SINR threshold for a single data rate on each card to simplify presentation. In Chapter 3, we mentioned that calibration is an important enabler for the SINR model to work. In Chapter 4, we show the calibration for each environment. Since in DIRC and Speed, we only need to choose the best antenna orientations that can maximize network capacity, and do not need to accurately compute the actual capacity, per-environment calibration is good enough for both systems to work well in practice. However, for power controlled systems, such coarse-grained calibration is not enough to choose the correct power levels, and we need

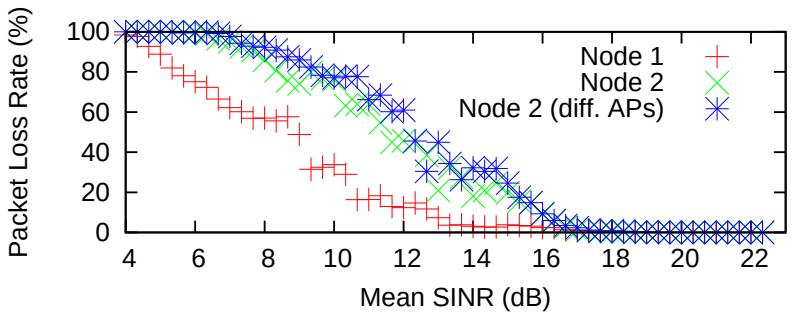


Figure 6.8: Calibration is needed

to further calibrate each card. Here notice that per-card calibration includes the effects of per-environment calibration, and per-card sensitivity. Figure 6.8 shows the SINR vs. loss rate curves for 24Mbps for two nodes (these two nodes have the leftmost and rightmost curves respectively). The difference between the lines is about 4dB. Based on this observation, instead of using a common SINR threshold for all wireless cards in our system, we calibrate the SINR thresholds for each wireless card, i.e., each card measures and broadcasts its own SINR thresholds. Next, we show how to calibrate the SINR threshold for each client.

The first step to calibrate the RSSI thresholds is to characterize the RSSI vs. loss rate curves at each client. This is done by having two APs (denoted as AP_1 and AP_2) send calibration frames simultaneously and having the client record each frame it receives and its RSSI reading. Each frame includes the total number of frames for this experiment and the ID of the current frame. The client C_i then calculates the mean RSSI and loss rates from AP_1 and AP_2 as $S_{1,i}$ and $S_{2,i}$, and $loss_{1,i}$ and $loss_{2,i}$, respectively. The loss rate calculation is based on the number of frames received from each AP and the total number of frames transmitted. Thus, client C_i can include two points to its SINR vs. loss rate curve, i.e., $(S_{1,i} - S_{2,i}, loss_{1,i})$ and $(S_{2,i} - S_{1,i}, loss_{2,i})$. Note that this requires client C_i to receive at least several frames from both APs. However, this requirement may not be satisfied if the signal from one AP is much stronger than the other. So in order to ensure that each client can receive at least several frames from both APs, we let the two APs to become unsynchronized for one timeslot, i.e., in the first timeslot, only the first AP is active and in the last timeslot, only the second AP is active, and in the middle timeslots, both are active. The loss rate calculation needs to be adjusted accordingly. Using this process, two points can be added to the SINR vs. loss rate curve at each client. In practice, usually only one of the two points (the one from with the higher SNR) provides valuable information. In order to fill in all the points along the curve, we repeat the above process but change the transmit power level on both APs. If more data points are needed, we change the pair of APs that are used for calibration. The calibration frames also include the transmit power level used for that frame and the AP that sends the frame. Figure 6.8 also shows the measured SINR vs. loss rate curves for the same node (node 2) with different calibrating AP pairs. The results show that the two curves are very close to each other. Similar observations can be made for most nodes in the network, with a few exceptions that we will discuss later. These observations indicate that 1) if necessary, measured curves from different AP pairs can be combined to form a whole curve, and 2) that the measured curve from one AP pair can be used when other APs act as senders or interferers. The

latter suggests that it is unnecessary to do per-pair calibration, and thus both the calibration and protocol overhead can be significantly reduced. In our experiments, we find that a one second experiment (which produces a pair of data points on the curve) is necessary for characterization to be accurate. Although it may incur high instantaneous overhead, this calibration is only needed once per environment, and can be recorded based on the AP the client associates with. Notice that it is possible to do this using only one AP, however, we find in our experiments that the calibration results are more accurate by having two APs. This may be due to the fact that there is external interference in the network and by explicitly introducing strong interference that can be counted in in the SINR model, the effects of external interference is removed. After this step, each client has fully sampled its SINR vs. loss rate curve.

The second step of calibration is to compute the offset between each client's SINR vs. loss rate curve and a reference curve. To simplify presentation, we first assume each SINR vs. loss rate curve has a single cut-off SINR threshold. Then the offset is the difference between the SINR threshold of the client and a reference threshold (i.e., 25dB). For more general SINR vs. loss rate curves, the offset is the number of shift from one curve to another. For example, in Figure 6.8, the offset between the SINR vs. loss rate curves from node 1 and node 2 is around 4dB, i.e. both curves will match if we shift the curve from node 2 to the left by 4dB. To compute the offset, we calculate the Euclidean distance between the two curve with various number of shift and choose the number that has the shortest distance. The offset with shortest distance is the calibration offset that each client broadcasts to the APs, and will be taken into consideration when APs estimate the link throughputs on clients.

Calibration Error

In our experiments, we observed that even when cards are calibrated in this way, the SINR vs. loss rate curves may still vary slightly depending on the multipath channel condition between the sender and receiver. These variations cause the power control protocol to perform poorly in certain scenarios and reduce throughput significantly on several links. In Opera, we take a conservative approach to deal with this problem by adding some value/slack to the SINR threshold / curve, which prevents frame collisions. In Opera, we set this value/slack to be 2dB, i.e., each threshold is increased by 2dB, based on the evaluation shown in Section 6.5.2.

6.4.2 Data Transmission

Since we leverage the directional MAC protocol proposed in Speed, the system operations for data transmission and channel reservation are very similar to that in Speed. The difference is that instead of running the antenna orientation algorithm, Opera APs run the power control algorithm. We will also highlight several parameters changes in Opera.

Next, we summarize Opera operations. When an AP has data frames to send, it reserves a timeslot by sending a *request-to-send* frame to the network. The key idea is that the AP needs to choose a power level that does not interfere with any existing reservations during the same timeslot (see Algorithm 7). Basically, each AP records all received *request-to-send* frames, and when an AP has traffic to send to a client, it runs Algorithm 7.

The *request-to-send* frame is a management frame and the AP uses the default CSMA approach to transmit these frames to avoid collisions. In order to ensure fairness among competing links, the window size for the *request-to-send* frames is saved across multiple timeslots. Thus, the link with earlier traffic arrival time has a higher chance of sending its *request-to-send* frame. When the frame has been transmitted over the air, the reservation has been confirmed, which means that all the other APs can transmit only when their transmissions do not interfere with this transmission. If the AP tries to send the *request-to-send* and before the frame can access the channel, the AP receives a *request-to-send* from another AP, the first AP will remove its own *request-to-send* from the transmission queue (txqueue) and run Algorithm 7 again.

To send data traffic simultanesouly, APs disable random backoff for data frames. This way all the APs will synchronize to transmit with each other after any data transmission (as there is no randomness), which effectively disables CSMA. Note that since link-layer ACKs ignore carrier sensing, it is possible to incur collisions on link-layer ACKs. However, as observed in our experiments and in previous work [109], the collisions rarely happen. Finally, note that all management frames are sent to a separate hardware txqueue with random backoff, i.e., where default CSMA is used. Also, all traffic from clients to APs, include TCP ACKs are transmitted with CSMA enabled, which can compete for the channel access right after they are generated.

6.4.3 Implementation

The Opera implementation is based on the Speed system implementation. We briefly discuss the implementation and highlight some changes in the system parameters below. We implemented Opera in Linux using Atheros 5213 cards. We work around several limitations of the driver, including lack of *request-to-send* withdrawal from the txqueue and making all data frames equal size.

There are several changes in system parameters, including, 1) 10ms timeslot size (instead of 20ms used by Speed), and even finer grained timeslot size is possible with a real-time operation system, which we left as future work, and 2) measurements are updated every one second (instead of 5 seconds), which is possible due to the lower measurement overhead of Opera over Speed (only 20ms measurement overhead for 4 APs, 2% overhead).

6.5 Evaluation

Our evaluation has two parts: a measurement-based evaluation of how well Opera and other solutions work in five different scenarios, and an end-to-end evaluation of the Opera implementation on a 10-node testbed.

6.5.1 Measurement-Based Evaluation

We use measurements collected in five different locations to evaluate the performance of Opera and other existing solutions. The goal of this part of our evaluation is to see how effectively the Opera algorithm optimizes network capacity. The solutions being evaluated are the same as presented in Chapter 6.2.1. The five locations include one research lab setting with 13 APs

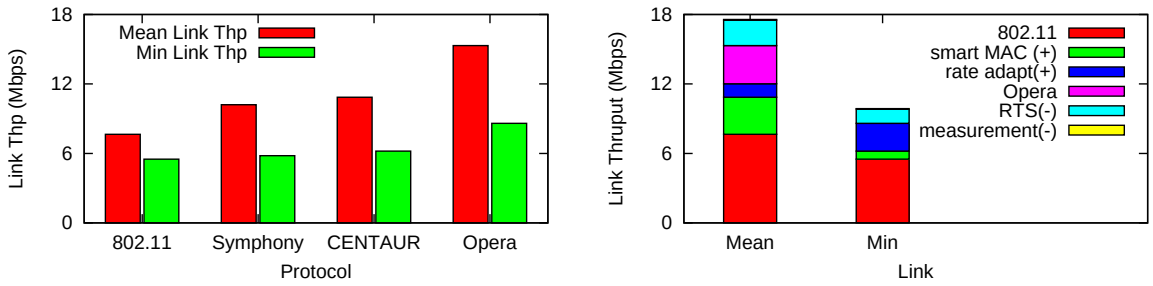
Scenario	# APs	# clients	AP ant.	client ant.
Campus 2	6	6	16,35°,omni	$4\times35°$ fan beam, omni
Lab 3	13	10	all avail.	omni
Campus 5	9	5	omni	omni
Campus 6	5	7	omni	omni
City	11	9	omni	omni

Table 6.3: Experimental setup for evaluating Opera

and 10 clients (Lab 3); three different campus settings with 6 APs and 6 clients (Campus 2), 9 APs and 5 clients (Campus 5), 5 APs and 7 clients (Campus 6); and one metropolitan area setting with 11 APs and 9 clients (City). In each setup, RSSI measurements from all APs to all clients are collected. In each experiment, we assume only 5 APs are active and we enumerate all possible combinations of 5 APs, and then present the mean and minimum link throughput from all AP combinations and all clients. The power level range of each sender is 20dB (from 0dBm to 20dBm).

Figure 6.4 shows the network capacity of different solutions in different locations. First, let us look at the Lab and the Campus 1 locations. Opera outperforms CENTAUR and Symphony by at least 40%. The performance of per-cell tuning and Opera is similar in the Campus 1 location. The reason is that since there are only 6 clients with 5 active APs, per-cell behavior is very close to per-link. In contrast, in the Lab scenario, Opera outperforms per-cell solution by 36% because there are 10 clients associating with 5 APs. Note that both scenarios are dense networks, i.e., most links can sustain high data rates when they are active one at a time, even when they use low txpower levels. In these cases, there is an optimal difference (in dB) of power levels (as illustrated in Section 6.4). The first node that reserves the channel uses a power level that is in the middle of the power level range, e.g., 10dBm. This leaves a large amount of slack in terms of tolerating interference for a subsequent sender. The second node that reserves the channel will set its power level accordingly, such that both transmissions can transmit simultaneously at high data rates. Thus, in these cases, the performance is very close to optimal.

For Campus 2, Opera outperforms CENTAUR and Symphony by 49%. In this scenario, the average AP-client distances are longer and thus the links can usually operate only at lower data rates, e.g., 24Mbps. The improvement comes from the second/third transmissions using lower data rates. Basically the first node that reserves the channel almost always use the highest power level so that it can support the highest data rate. When the second node reserves the channel, it uses a low power level and a low data rate, such that it does not affect the first reservation. In this scenario, the optimal solution achieves much higher capacity because it considers all links at the same time, e.g., both transmissions may use lower data rates and do not interfere with each other. However, the optimal solution requires a centralized controller and does not apply to our distributed setting. Another observation is that even though there are only 5 clients in the network, Opera still outperforms the per-cell protocol by 25%. This improvement comes from the fact that the choice of txpower levels in Opera is also fine-grained in time. By selecting



(a) UDP (b) UDP Breakdown

Figure 6.9: UDP performance of Opera

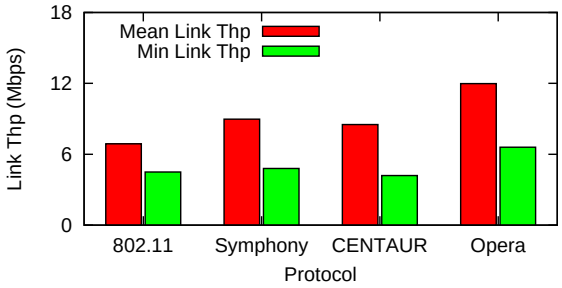


Figure 6.10: TCP performance of Opera

power levels for each timeslot, each link has the flexibility to choose the most appropriate power levels according to the interfering links. In contrast, the per-cell solution will always use a fixed txpower level that maybe good for one interferer but bad for another.

For Campus 5, the benefit of Opera is relatively low (14% over existing solutions). From inspecting the signal strength measurements, we can identify that Opera provides less benefit when several clients are closely located in signal space (the signal strength from the clients to multiple APs share similar patterns).

In fact, the performance of the per-link transmit power control and CCA tuning is reasonable: the performance is close to that of Opera except in one case, where it is 35% worse than Opera. But as explained in Chapter 6.3.2, this approach is still undesirable. In summary, the results show that our heuristic algorithm can effectively improve network capacity in various scenarios over existing solutions.

6.5.2 End-to-End Evaluation

In this section, we evaluate Opera on a 10-node indoor testbed. The locations of the 10 nodes are shown as $n_1 - n_{10}$ in Figure 5.8(c), i.e., we only use the omni-directional nodes in the Lab 3 scenario.

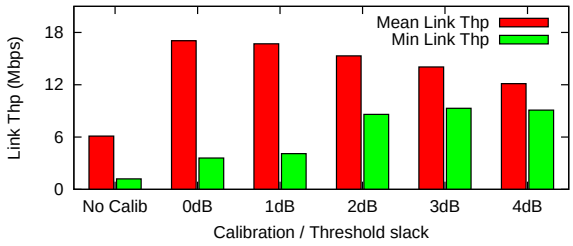


Figure 6.11: Evaluation of Opera's calibration mecahnism

Network Capacity

First, we evaluate the UDP and TCP capacity of Opera on the testbed. To evaluate our system in multiple dense network configurations, we use the 10 nodes and enumerate all possible combinations of 4 nodes as APs (and 6 others as clients). We associate each client with the closest AP, and if there is an idle AP where no clients associate with that AP, we skip that combination.

The approaches we tested are: the default 802.11 configuration, Symphony, CENTAUR, and Opera. In the UDP test, all APs transmit simultaneously 30Mbps of UDP traffic for 10 seconds. In the TCP test, we perform a TCP bulk transfer from the APs to all associated clients. Note that for Opera, the wireless data rate is chosen by the algorithm and rate adaptation is enabled.

The minimum and mean link throughputs of all six flows across all AP combinations are shown in Figure 6.9 and Figure 6.10. The results show that: 1) Opera improves network capacity over other protocols by 40% for both UDP and TCP, and 2) the minimum link throughput of Opera is also higher than that of other protocols. In fact, Opera improves throughput on all links. This illustrates that by optimizing spatial reuse, Opera significantly improves all link throughputs and, thus, network capacity over existing solutions.

Figure 6.9(b) shows the breakdown of protocol gain (+) and overhead (-), for both mean and minimum link throughputs for UDP. It shows that for mean link throughput, 20% of the gain comes from a better MAC protocol, 8% gain comes from the fact that data rate adaptation is enabled, 22% gain comes from the fine-grained power control algorithm, and the overhead of *request-to-send* messages is about 14%, and 2% protocol overhead for collecting measurements.

Calibration

Next, we evaluate how well the calibration mechanism presented in Chapter 6.4 works. Figure 6.11 shows the mean and minimum link throughputs with and without calibration (No Calib vs. 0dB). The results show that without calibration, the network capacity is very low due to packet collisions between nodes that should not have been transmitting concurrently. By enabling calibration, the capacity is much improved. However, the minimum link throughput is still low due to calibration errors. As mentioned in Chapter 6.4, we use SINR threshold/curve slack to account for the inaccuracies in SINR model and calibration. Generally speaking, a large slack can account for more errors but at the same time may prevent non-interfering transmissions from happening simultaneously. This is shown in Figure 6.11, where we tested different SINR threshold slacks (0-4dB). By putting more slack in SINR threshold, the mean throughput drops

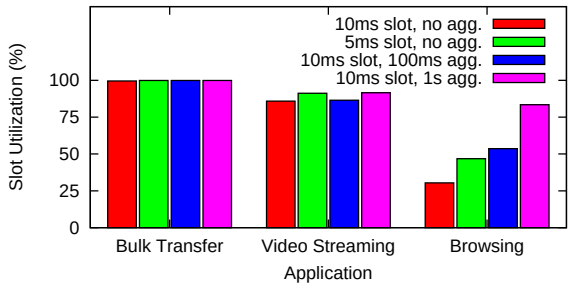


Figure 6.12: Timeslot utilization under real traffic

and minimum throughput increases until 3dB. In our system, we choose the slack to be 2dB since it best balances between mean and minimum link throughputs.

Timeslot Size and Traffic Patterns

In Opera, we set the timeslot size to be 10ms, which roughly equals to 45KB of data with 54Mbps data rate. This indicates that, in the worst case, when the data to be transmitted in the timeslot is less than 45KB, Opera may be wasting the rest of the timeslot. This is especially high overhead for short TCP transfers and during TCP slow start. Here, we demonstrates this effect and discuss how to deal with this problem. Figure 6.12 shows the slot utilization (excluding reservation overhead) under three applications with different traffic patterns. Note that experiments in this section are carried out with a wireless client accessing services from remote servers in the Internet (around 80ms RTT). The result suggests that without any modifications, the default Opera system works fine for bulk transfer and video streaming, but does not work very well for web browsing with small TCP transfers.

There are several approaches to deal with this problem: 1) Shrink the size of the timeslot. (e.g., see the 5ms slot, no agg bars in Figure 6.12) This solution does not completely solve the problem because the utilization for web browsing remains low. Also, by reducing the timeslot size, the percentage overhead of channel reservations also increases, unless we allow APs to reserve multiple timeslots in one reservation message. 2) Have APs aggregate data traffic over a period of time to a single client (e.g., see the 10ms slot, 100ms/1s agg bars in Figure 6.12). This is commonly done in existing power management protocols. Note that only traffic from non-delay-sensitive applications, e.g., low priority traffic in 802.11e, should be aggregated (since the latency can be high). 3) Allow the APs to send data frames if the channel is clear for some time, indicating that all transmissions have finished. 4) Allow the APs to explicitly cancel a reservation so that other APs can make reservations for the rest of the timeslot. 5) Have APs aggregate data frames for multiple clients in a single timeslot. In practice, an Opera deployment could use a combination of these solutions.

Latency

Since Opera's MAC protocol is based on timeslots, the latency of certain packets can be very large, a node will have to wait for at least one timeslot before it can transmit again. Here, we

Protocol	Mean Lat. (ms)	99th PCTL Lat. (ms)
802.11	1.8	3.4
Opera	0.8	41.2

Table 6.4: Protocol latency

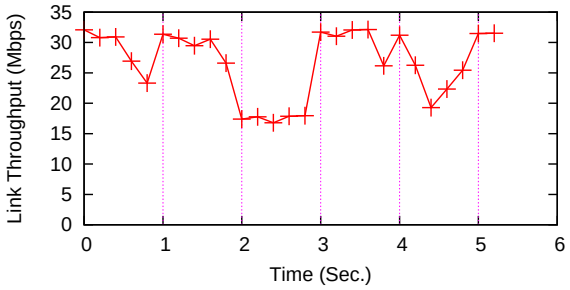


Figure 6.13: Handling mobility in Opera

evaluate the latency properties of Opera and 802.11. We define latency as the time difference between two consecutive frames are received. Table 6.4 shows the mean and 99th percentile latency across all consecutive frames and all links. It shows that while the mean latency of Opera is much reduced over CSMA based 802.11, the 99th percentile latency of Opera is much higher. Note that we assume Opera APs have some knowledge on the application and do not use traffic aggregation for delay sensitive applications. Note that the latency introduced is still tolerable for most multimedia applications (note that there is very low latency for traffic from clients to APs since they use CSMA). For highly delay sensitive applications, we recommend using even finer grained timeslots to reduce worst case latency and at the same time gain benefits over existing solutions.

Mobility

In this experiment, we evaluate how well Opera handles client mobility. We enable two transmissions in the testbed, i.e., two APs transmitting to two clients, and move one of the two clients around the room. Figure 6.13 shows the link throughput for the moving client. The results suggest that by updating measurements every one second (the measurement process takes 20ms), Opera can handle node mobility very well at walking speed. For reference, two 802.11 transmissions will achieve throughput around 17-18Mbps.

6.6 Hardness of the Problem

As explained in Section 6.3.4, to enable any two concurrent transmissions, their power levels P_1 and P_2 need to be set such that

$$SINR_1 = P_1 - L_{11} - P_2 + L_{21} \geq SINR_{thrsh}$$

and

$$SINR_2 = P_2 - L_{22} - P_1 + L_{12} \geq SINR_{thrsh}$$

That is,

$$L_{12} - SINR_{thrsh} - L_{22} \geq P_1 - P_2 \geq SINR_{thrsh} + L_{11} - L_{21}$$

Thus for a set of transmissions, the problem of maximizing spatial reuse can be formulated as finding their power levels, i.e., a set $\{P_i\}$ such that satisfies the largest number of inequalities of

$$c_{ij} \geq P_i - P_j \geq d_{ij}$$

We call the above problem Max-Ineq defined as follows.

Definition 6.6.1 *A Max-Ineq instance I is defined by a set of variable $P_i \in Z$ and a set of inequality over these constraint of the form $c_{ij} \geq P_i - P_j \geq d_{ij}$. The algorithmic task is to find the assignment of P_i to maximize the number of satisfied constraints. We use $opt(I)$ to be the maximum number of constraints that can be satisfied by any assignment.*

We claim there is no $O(1)$-approximation algorithm for this problem assuming the Unique Games Conjecture [62], defined as follows.

Definition 6.6.2 *A unique game conjecture is defined by a set of variable $P_i \in [q]$ and a set of constraints of the form $P_i - P_j = c_{ij}(mod\ q)$. The algorithm task is to find some assignment of P_i to maximize the number of satisfied constraints. For a given Unique Games instance U, we use $opt(U)$ denote the maximum number of constraints that can be satisfied by any assignment of the variable P_i.*

Conjecture 6.6.3 *(Unique Games Conjecture) For any ϵ, there exists some big enough q such that it is NP hard to distinguish the following two two cases: i) $opt(U) \geq 1 - \epsilon$ ii) $opt(U) \leq \epsilon$.*

Now we prove there is no $O(1)$-approximation algorithm for Max-Ineq. Following is the reduction from the Unique Games Conjecture.

Then for any constraint Unique Games Instance U,

$$P_i - P_j = a_{ij}(\mathrm{mod} q) \tag{6.1}$$

we can be rewritten as two equations

$$a_{ij} \geq P_i - P_j \geq a_{ij}$$

$$a_{ij} - q \geq P_i - P_j \geq\ a_{ij} - q$$

This becomes a corresponding Max-Ineq problem I. Thus Max-Ineq has the following two properties:

1. if $opt(U) \geq 1 - \epsilon$, then $opt(I) \geq 1/2 - \epsilon/2$.
2. if $opt(U) \leq \epsilon$, then $opt(I) \leq \epsilon/2$.

By the unique games conjecture, for any Max-Ineq instance I, it is NP-hard to distinguish whether $opt(I) \geq 1/2 - \epsilon/2$ and $opt(I) < \epsilon/2$. This implies there is no constant approximation for the Max-Ineq problem.

It is also worth noting that the problem is conditional NP-hard only if no assumptions are made on the path losses. In fact, if free space path loss model is assumed, the problem of finding optimal power levels can be solved in polynomial time [38].

6.7 Summary

In this Chapter, we presented the design, implementation, and evaluation of Opera, a distributed and fine-grained power control protocol that optimizes spatial reuse in omni-directional networks. Similar to Speed (Chapter 5) and DIRC (Chapter 4), Opera relies on the SINR model. Also, Opera uses Speed's distributed MAC protocol to coordinate transmissions.

First, we show that existing solutions do not work well in chaotic networks where node densities are highly variable and interference is unpredictable. Opera addresses several challenges, both algorithmical and practical. Our measurements of Opera in five different locations and evaluation of Opera implementation on a 10-node testbed show that Opera works well in practice.

Chapter 7

The Separation Metric

In Chapter 2, Chapter 4, and Chapter 5, we showed that directional antenna systems can be very effective in two indoor experimental settings. Related work [20] also made similar observations. However, it has also been observed that spatial reuse can only be slightly improved in the testbed presented in [69]. These contradicting conclusions are caused by the difference in experimental settings. In the first part of this chapter, we discuss three major factors that determine the effectiveness of directional transmission and power control: the workload, the interference conditions, and the multipath channel conditions. Depending on the particular environment, different conclusions can be drawn with regard to whether directional transmission and power control are useful in that particular environment. In this case, the contradicting conclusions are caused by different workload (or more specifically, node density) between the environments used in these studies. Since the testbed used in [69] is a sparse network with low node density, the level of contention in the testbed is low. Thus the benefits from stronger signal strength dominate over those from spatial reuse.

Consequently, to estimate the benefits of converting a subset of APs and/or clients to using directional antennas in a particular environment, it is crucial to estimate the three major factors. However, it is very challenging to estimate the third factor of the multipath channel conditions since multipath have both good and bad effects on spatial reuse. As a result, it is challenging to determine the effectiveness of directional transmission and power control in any particular environment. Specifically, given any particular environment, we would like to answer the following three key questions.

- Is it worth deploying directional antennas in a particular environment? Or is it worth enabling power control in a particular environment?
- What type of antenna capabilities (e.g. directional or omni-directional and beam-width) should each AP and client have?
- What are the good locations for directional APs in the deployment?

In the second part of this chapter, we introduce the *separation metric* [72] to answer these questions. The idea of the separation metric is to capture both the angular and the distance separation for pairs of transmissions in a way with low measurement and computational costs. This way the separation metric can capture the various properties of the deployment. Our experimental results show in several scenarios that the separation metric has a positive correlation

with the network capacity, i.e., a higher separation metric indicates a higher network capacity and vice versa. Due to the positive and reasonably strong correlation of the separation metric and the network capacity, we can estimate the effectiveness of directional transmission or power control, and guide the placement of directional APs in any particular environment. In this part of the chapter, we assume a single administrative domain, such as enterprise networks, campus networks, or other wireless service providers, where network planning is feasible.

7.1 Three Major Factors

In this section, we examine three major factors that affect the effectiveness of directional transmission and power control and the benefits of our proposed solutions over naive (i.e., uncoordinated) solutions. The three factors are the workload, the interference conditions, and the multipath channel conditions in the particular environment.

7.1.1 Factor I: Workload

As we have discussed in Chapter 1, directional transmission and power control are quite useful when the workload in the network is high. In this case, enabling spatial reuse that allows multiple senders to transmit simultaneously becomes critical. In application scenarios with low workload, power control is useful only to reduce power consumption, and directional transmission can still be beneficial since it can increase the received signal strength and thus the transmission rate. In this case, the uncoordinated solutions discussed in Chapter 1.1.1 suffice, and our proposed solutions can only slightly improve the system performance, as observed in previous work [69]. The workload in a network is determined by node density and traffic patterns.

- Node density: Node density is the number of interfering nodes within a certain area. When the node density is high, the expected workload in the network is high. When the node density is low, the expected workload is low. Here we would like to note that when the node density is extremely high (100s of nodes), our current systems would incur very high overhead in taking all the measurements due to the large number of neighboring APs. One possible way to deal with this is to reduce the number of neighboring APs by reducing the power levels on the senders.
- Traffic patterns: When the applications run on the interfering nodes require high bandwidth, the expected workload is high. When all the applications only produce bursty traffic, the expected workload is low.

7.1.2 Factor II: Interference Conditions

The second factor is the interference conditions. While the first factor of workload affects the level of interference or the level of contention in the networks, in this section, we primarily discuss the physical layer interference properties. Similar to the first factor, when the interference in a network is high, the expected benefits from directional transmission and power control is high. The interference conditions is determined by the following factors:

- Signal propagation range: When the signal propagation range is long, the expected level of contention is higher, because each sender is expected to contend with more other senders. When the propagation range is short, the expected level of contention is low.
- Effects of obstacles: When the signal strength decreases a lot while passing through obstacles, the expected level of contention is low, because interference is confined inside each room and the senders in one room may not contend with the senders in another room. When signals do not degrade much passing through obstacles, the expected level of contention is high.
- Antenna beamwidth: When the antenna beamwidth is narrow, the expected level of contention is low, since senders contend with fewer other senders. When the antenna beamwidth is wide, the expected level of contention is high. This applies to antenna beamwidth on both the APs and the clients.

7.1.3 Factor III: Multipath channel conditions

The third factor is the multipath channel conditions in a particular environment. Generally speaking, with higher degree of multipath, the effectiveness of directional antennas will be reduced, because the level of directionality is decreased [20]. At the same time, the benefits of our approach over naive solutions will increase. The multipath channel conditions are primarily determined by the particular environment and the locations of the nodes. For example, generally speaking, the degree of multipath is lower in outdoor space and higher in indoor environments.

In contrast to the previous two factors, it is very difficult to determine the degree of multipath, especially in rich-scattered indoor environments, due to the lack of a reasonable indoor signal propagation model. Later in this chapter, we present the separation metric, which is a measurement based metric that characterizes the benefits of directional transmission and power control in a particular environment, by capturing the second and the third factors.

7.2 Case Study

We presented three factors that affect the benefits of directional transmission and power control. Based on the factors, in this section, we discuss the applicability of our conclusions to various other environments, especially the scenarios that are different from our experimental setups in the first and the second factors. We will focus on the third factor in the second part of this chapter.

7.2.1 Existing Wireless Networks

Most existing wireless networks operate in 2.4GHz. Thus the signal propagation range, the effects of obstacles, and the antenna beamwidth are similar to that in our experiments. Now we look at other factors.

- Node density: The node density in our experiment setup is high. For example, in most setups, we put 6 APs and 6 clients within a 20 m$\times$20 m area. The node density in existing

wireless networks varies from scenarios to scenarios. For most enterprise and campus networks, the node density is low since all the AP locations are planned and the primary goal of the planning is to ensure coverage. Thus the level of contention is low in these scenarios. The node density in metropolitan areas and in apartment buildings is significantly higher, similar to that in our experimental setup.

- Traffic patterns: The traffic pattern in our experiments is significantly different from that in existing networks. In our experiments, we primarily focus on saturated traffic. While in existing networks, most applications produce bursty traffic. Thus the level of contention in existing networks is expected to be much lower than that in our experiments. However, this is also partly caused by the fact that the performance of existing wireless networks is not good enough for users to comfortably shift their applications that require high bandwidth to wireless. With the availability of multi-antenna systems such as MIMO and directional antennas, new wireless applications that require high bandwidth, such as wireless hubs and wireless displays, are emerging.
- Multipath channel conditions: The multipath channel conditions vary from scenarios to scenarios and we try to capture these channel conditions by carrying out experiments in two very typical but different indoor testbeds. Since the application scenarios targeted in our work are indoor environments, we did not do extensive experiments in outdoor scenarios.

In summary, the workload of existing wireless networks is much lower than that of our experimental setup. This indicates that our proposed solutions are less effective in existing wireless networks. However, we believe that with the emerging wireless technologies and applications, the workload (both the node density and the amount of traffic) will keep increasing.

7.2.2 White Spaces

These are the Ultra High Frequency (UHF) bands that will be used in future White Space wireless networking, especially between 510 and 700 MHz. Next we compare this frequency band with 2.4GHz.

- Signal propagation range: The propagation range in UHF bands is very far, which will increase the level of contention.
- Effects of obstacles: Signals do not degrade much in these bands, thus the level of contention is increased.
- Antenna beamwidth: Since the size of directional antennas in these bands is inherently large, the beamwidth on clients is expected to be wide. Thus the level of contention is increased.
- Node density: Due to the large antenna size and the long propagation range, the usage model in these UHF bands is likely to be similar to that in cellular networks, i.e., service providers plan the locations of the APs, and each AP service a particular area. Thus node density (or AP density) is expected to be lower than that in 2.4GHz.
- Traffic patterns: Since each AP services multiple clients, the aggregated amount of traffic is expected to be high, and thus the workload is expected to be high.

- Multipath channel conditions: The degree of multipath is expected to be lower since the application scenarios are mostly outdoor environments.

In summary, both the workload and the level of contention are expected to be higher in UHF than in 2.4GHz, indicating that directional transmission and/or power control will be more useful than those in 2.4GHz. Also, our proposed systems are expected to be less beneficial over the naive solutions in these bands.

7.2.3 60GHz (Millimeter Bands)

Recently researchers have started to exploit the largely unused 60GHz bands [6]. Next we compare this frequency band with 2.4GHz.

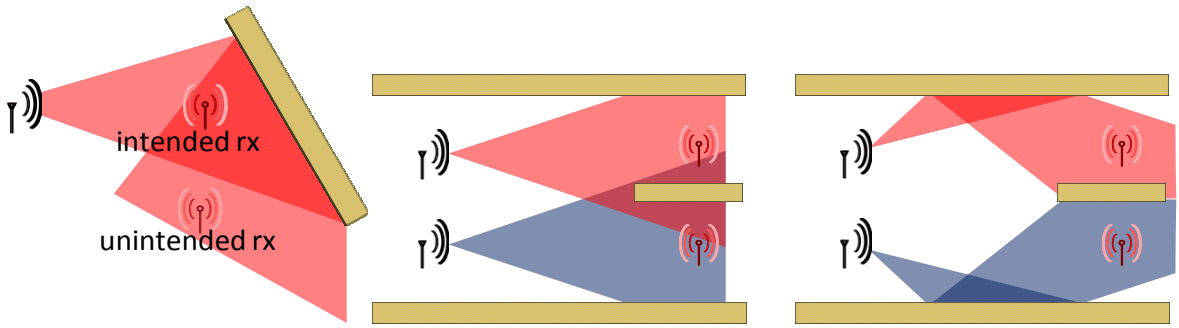
- Signal propagation range: In these extremely high frequency bands (or millimeter bands), signals are very prone to atmospheric attenuation and thus the propagation range is short. This indicates that the level of contention is expected to be low. Also due to atmospheric attenuation, only directional antennas can deliver the signals to the receiver with reasonable signal strength.
- Effects of obstacles: Signals degrade much passing through obstacles in these bands, thus the level of contention is decreased.
- Antenna beamwidth: The short wavelength makes it possible to manufacture very small antennas with very narrow beamwidth, on both the AP side and the client side.
- Node density: It has been envisioned that hundreds of devices will operate in these bands within a small area, indicating high workload.
- Traffic patterns: While most applications in 60GHz only use the network lightly, some applications use the network very heavily, such as high definition video streaming.
- Multipath channel conditions: The degree of multipath is expected to be lower since most signals are confined inside a room. However, experiments have shown that signals can still reach destinations through reflectors [6], indicating that the degree of multipath can still be high even inside a room.

Directional antennas are critical in 60GHz wireless transmission, and even with very narrow beams, interference can still be a problem due to the high node density, indicating that our solutions are useful in 60GHz bands.

In summary, we believe that both directional transmission and power control will still be very useful in dense networks that operate in frequency bands other than 2.4GHz, especially in UHF and 60GHz bands. However, further studies are needed to quantify the actual benefits of both techniques in these bands.

7.3 Challenges and Overview of The Separation Metric

In the previous sections, we presented three factors that affect the benefits of directional transmission and power control and showed how to roughly estimate the first two factors based on the descriptions of the application scenario and the environment. However, the third factor of



(a) Downside (cause unintended interference)

(b) Upside (alternative path to improve spatial reuse)

Figure 7.1: Downside and upside of rich scattering

the multipath channel conditions is much harder to estimate. This is especially true in rich-scattered indoor environments, since the rich scattering (or multipath) can have both the upside and the downside on directional antenna systems. Figure 7.1 shows an example. On the downside (Figure 7.1(a)), rich scattering may cause unintended interference, i.e., strong interference at the unintended receivers through reflectors. On the upside (Figure 7.1(b)), alternative paths and the obstacles make it possible to reduce interference between transmissions and allow both transmissions to occur concurrently. The above statements are also true for power-controlled systems.

Consequently, it is challenging to determine the effectiveness of directional transmission and power control in any particular environment. Throughout this dissertation, we have used two approaches to answer this question, 1) direct measurements of network capacity, and 2) simulation based on SINR model. Both approaches suffer from scalability problem either in terms of the number of measurements or the computational cost.

The first naive solution is to directly measure the network capacity that can be achieved in each particular environment, with one measurement for each possible AP placement. This naive approach involves exhaustive measurements and is thus not practical. Furthermore, this approach involves deploying multiple directional APs before knowing whether they will be effective.

The second solution is to carry out a site survey that measures the signal strength from all the APs to all the clients, and then use the SINR model to compute the network capacity. The site survey process involves moving one (directional) AP across all candidate locations with all the clients located at their locations, and collecting signal strength measurements from that AP location to all the client locations. The site survey process can be automated by using a robot [63]. This is the approach we take in our measurement studies throughout this dissertation. Using this approach, the measurement overhead is significantly reduced, but the computational cost is still high, i.e., exponential in the number of APs in that scenario. In Chapter 2.2.3, we presented how we compute the network capacity over a collection of APs and clients for both directional transmission and power control, i.e., using the algorithms of MaxCAP, OptPC, and OptMAC. For example, for directional transmission, OptMAC does an exhaustive search on all possible schedules, and the number of possible schedules is $\binom{N}{1} + \binom{N}{2} + \ldots + \binom{N}{N}$. And for the schedule of i transmissions, MaxCAP does an exhaustive search on all possible antenna orientations, thus

the complexity for that schedule is $(K_{AP} * K_C)^i$. In summary, the complexity to compute the network capacity is $(K_{AP} * K_C)\binom{N}{1} + (K_{AP} * K_C)^2\binom{N}{2} + \ldots + (K_{AP} * K_C)^N\binom{N}{N}$. In fact, due to the complexity, we can only compute the network capacity for at most 6 APs using this approach in our experiments.

Another naive solution that has been used in current practices is simply to maximize the coverage, or equivalently to maximize the minimum signal strength across all the clients. We will show later in this chapter that this approach cannot maximize spatial reuse.

We take one step further and define the separation metric, which abstracts out the most important aspects that enable spatial reuse and which at the same time makes the computation manageable (the computational cost is polynomial to the number of APs and clients). Unavoidably, the simplifications to reduce the complexity will also reduce the accuracy of the separation metric in predicting the network capacity. In Chapter 7.5, we discuss two simplifications and their implications. We also show that even with the simplifications, the separation metric still has a positive and reasonably strong correlation with the network capacity in several experimental scenarios, especially for directional transmission.

There are two steps in computing the separation metric. The first step is the site survey process that characterizes the signal strength from the APs to the clients, as described above. The second step is to compute the separation metric of pairs of transmissions based on the signal strength measurements.

After computing the separation metric, it can be used for the following tasks.

- By comparing the separation metric of different directional antenna configurations, we can estimate and compare the performance (network capacity) of different directional antenna configurations and estimate the performance (network capacity) benefits of power control. This is presented in Chapter 7.6.1 and Chapter 7.6.2.
- By carrying out measurements or site surveys with different antennas (e.g., with different beamwidths or antenna patterns), we can estimate the performance benefits (network capacity) of these different antennas. This is presented in Chapter 7.6.1.
- By comparing the separation metric at different AP locations, we can estimate the best locations to place the APs in a particular environment that maximize network capacity or spatial reuse. This is presented in Chapter 7.6.3.

There are several variants of the separation metric definition, based on 1) whether it is computed from the point of view of an AP, a client, or a link, and 2) whether client-AP association is considered. These different variants have different computational costs and can achieve different levels of accuracy. Generally speaking, we find that the separation metric defined on link pairs while considering association is the most accurate but, at the same time, incurs the highest computational cost (still polynomial).

Now we discuss the scope of the separation metric. Similar to the optimization goal of our proposed systems, the metric to evaluate the separation metric is the network capacity or spatial reuse. This is in contrast to previous work on wireless network planning that only focuses on wireless coverage. Also, we have identified three factors that would affect network capacity, thus ideally we would like the separation metric to capture all three factors. However, the separation metric defined in this chapter does not capture the first factor of workload (or equivalently, node density and traffic patterns); we simply assume dense network deployment and saturated traffic

patterns. The reason is that as explained before, the degree of spatial reuse for sparse networks or bursty traffic loads will be low, and thus we can conclude that power control is not very useful and directional antennas are useful only for increasing signal strength, without taking any measurements. This also indicates that the separation metric cannot be applied to sparse networks or bursty traffic loads. Finally in order to calculate the separation metric, we assume that the locations of the clients are known. Later we will show in this chapter that exact locations of the clients are unnecessary and some perturbations (1-2m) in client locations are acceptable.

Note that the separation metric cannot be used to predict the actual network capacity in a particular setup because the separation metric depends on the location of the clients, and thus can only be compared within the same environment (e.g., directional vs. omni-directional antennas in the same setup), but not across different environments (e.g., directional antennas in one setup vs. another).

7.4 Separation Metric

In this section, we present all variants of the separation metric for directional antennas and power control.

7.4.1 Directional Antennas

The separation metric for directional antenna systems has six variants, 1) it can be defined for APs, clients, or link pairs, and 2) each has an association-aware and an association-ignorant version.

AP Separation Metric

First, we provide some intuition to motivate the separation metric. Given a single AP (AP_i), an intended receiver C_{j_1} and an unintended receiver C_{j_2}, the potential for spatially reusing spectrum (or "goodness") can be characterized by how well the AP can isolate the two clients, i.e., can the AP use some orientation to transmit to the intended client without causing interference at the unintended client. The below separation metric captures this:

$$SEP_{AP}(i, j_1, j_2) = \max_{k_i \in K_{AP}, k_{j_1}, k_{j_2} \in K_C} S(i, j_1, k_i, k_{j_1}) - S(i, j_2, k_i, k_{j_2})$$

This metric basically shows the maximum difference between the signal strength from AP_i to an intended receiver C_{j_1} and to an unintended receiver C_{j_2}, across all antenna orientations on both the AP and the two clients. High separation metrics (> 25dB) make it more likely that the two clients, served by two different APs can successfully receive frames at the same time since the SINR threshold for receiving packets successfully at 54Mbps is 25dB.

The overall "goodness" of an AP, (AP_i), can be obtained by aggregating the separation of all client pairs:

$$SEP_{AP}(i) = \sum_{j_1, j_2} SEP_{AP}(i, j_1, j_2)$$

For a network of APs and clients, the "goodness" of that network can be obtained by aggregating the separation from all APs:

$$SEP_{AP} = \sum_{i} SEP_{AP}(i)$$

Note that the AP separation metric can be used to visualize the "goodness" of each location. As we show later, it can also be used to decide where to place directional antennas in an incremental fashion, especially when the candidate AP locations are sparse (e.g., when upgrading an existing omni-directional antenna setup). The problem with this metric is that it only considers a single AP (the signal and interference level from this AP to the clients) and ignores the interactions between the APs (the actual SINR is computed with multiple APs). Thus it does not work very well when candidate AP locations are dense. For example, placing multiple APs at the same location with the highest separation metric will maximize the total separation for that network, but not necessarily provide the best performance. In order to deal with this issue, we use other variants of the separation metric, which considers interactions among the APs.

Client Separation Metric

For a client C_j, a transmitting AP_{i_1}, and an interfering AP_{i_2}, the separation metric is defined as:

$$SEP_C(i_1, i_2, j) = \max_{k_{i_1}, k_{i_2} \in K_{AP}, k_j \in K_C} S(i_1, j, k_{i_1}, k_j) - S(i_2, j, k_{i_2}, k_j)$$

The client separation metric characterizes the ability of the client to receive from its sender while shielding from the interference source. The client separation metric can be aggregated on each client:

$$SEP_C(j) = \sum_{i_1, i_2} SEP_C(i_1, i_2, j)$$

For a collection of APs and clients:

$$SEP_C = \sum_{j} SEP_C(j)$$

Link Pair Separation Metric

For a pair of links $AP_{i_1} \rightarrow C_{j_1}$ and $AP_{i_2} \rightarrow C_{j_2}$, the separation metric is defined as:

$$\begin{aligned} SEP_P(i_1, i_2, j_1, j_2) = \max_{k_{i_1}, k_{i_2} \in K_{AP}, k_{j_1}, k_{j_2} \in K_C} & \quad S(i_1, j_1, k_{i_1}, k_{j_1}) - S(i_2, j_1, k_{i_2}, k_{j_1}) \\ + & \quad S(i_2, j_2, k_{i_2}, k_{j_2}) - S(i_1, j_2, k_{i_1}, k_{j_2}) \end{aligned}$$

The separation metric for the link pair is essentially the sum of SINR on both links. The link pair separation metric can be aggregated on a collection of APs and clients:

$$SEP_P = \sum_{i_1, j_1, i_2, j_2} SEP_P(i_1, i_2, j_1, j_2)$$

Separation Metric with Association

The previous definitions of the separation metric do not consider AP-client association, which is an important factor in determining network capacity. So we also introduce the separation metric with association, here we associate each client with the AP with the strongest signal strength. We use $ASSOC(AP_i)$ to denote the clients that associate with AP_i. Then the association-aware AP separation metric becomes:

$$SEP_{AP,ASSOC}(i) = \sum_{j_1 \in ASSOC(AP_i), j_2 \notin ASSOC(AP_i)} SEP_{AP}(i, j_1, j_2)$$

$$SEP_{AP,ASSOC} = \sum_i SEP_{AP,ASSOC}(i)$$

The client separation metric becomes:

$$SEP_{C,ASSOC}(j) = \sum_{j \in ASSOC(AP_{i_1}), j \notin ASSOC(AP_{i_2})} SEP_C(i_1, i_2, j)$$

$$SEP_{C,ASSOC} = \sum_j SEP_{C,ASSOC}(j)$$

The link pair separation metric becomes:

$$SEP_{P,ASSOC} = \sum_{j_1 \in ASSOC(AP_{i_1}), j_2 \in ASSOC(AP_{i_2})} SEP_P(i_1, i_2, j_1, j_2)$$

Where Does Separation Comes From?

Next, we examine the two components that contribute to the separation: distance and angular separation. The separation metric, in fact, is the sum of these two components.

Distance separation can be achieved if the AP is closer to its intended clients than the unintended clients in signal space, and the larger the difference, the higher the distance separation. Thus, the distance separation is mostly affected by the AP location, client locations, and the RF environment. Note that the notions of "far" and "close" apply only in signal space, and may not apply in physical space. For example, we find zero correlation in the physical space and the signal space in the Campus 3 scenario, i.e., no correlation between the actual signal strength and the predicted signal strength using the free space loss model.

Angular separation can be achieved using directional antennas on APs and/or clients. The contribution of angular separation comes from antenna orientations that allow the AP to avoid interfering the unintended clients while maximizes the signal strength at the intended clients, even without distance separation. Note that the angular separation only applies to directional antennas, i.e., angular separation for omni-directional antennas will be 0. Thus, angular separation is determined by the degree of directionality (antenna beamwidth), the RF environment, and the locations of the APs and clients. Narrower beam antennas will provide higher angular separation. Also, generally speaking, the angular separation is high if the AP is in the middle of the clients in the signal space. This is because different clients show up in different sectors on

the AP. For example, when the AP is transmitting to a client to the left of the AP, it incurs low interference level on the clients to the right of the AP. In contrast, angular separation is low if the AP is in the corner of the deployment and all clients show up in a small subset of the sectors of the AP. Note that again, the notions of "middle" and "corner" apply only in signal space.

7.4.2 Power Control

For power control, the separation metric is only defined on link pairs. As presented in Chapter 6, for two flows, power control essentially redistributes the sum of SINR values between the two links. This indicates that with power control, the sum of SINR values can be even distributed across the two links and can be used efficiently this way. On the other hand, without power control, the sum of SINR values will be the same, but one may have a very high SINR and the other may have a low SINR. However, since for 802.11 systems, SINR values above a threshold are not useful (the highest data rate is 54Mbps), this uneven allocation may be inefficient. For example, if the sum of SINR for two flows is 50dB, then power control can allow both transmissions to occur simultaneously without interfering. Without power control, however, it could be that the SINR for the first flow is 30dB (but it only needs 25dB, so waste 5dB) and the SINR for the second flow is 20dB. In fact, this also illustrates that as faster data rates are defined above 54Mbps, power control becomes less useful. Given the above intuition, we define the separation metric for no power control. Note that the separation metric for power control is only defined over link pairs, since that defined over APs and clients makes much less sense. We first define a utility function $ceiling$ as:

$$ceiling(SINR, SINR_{thresh}) = \begin{cases} SINR_{thresh} & \text{if } SINR > SINR_{high} \\ 0 & \text{if } SINR < 0 \\ SINR & \text{otherwise} \end{cases}$$

Thus for a pair of links $AP_{i_1} \rightarrow C_{j_1}$ and $AP_{i_2} \rightarrow C_{j_2}$, the separation metric defined for no power control is:

$$\begin{aligned} SEP_{NOPC}(i_1, i_2, j_1, j_2) = \; & ceiling(S(i_1, j_1, k_{i_1}, k_{j_1}) - S(i_2, j_1, k_{i_2}, k_{j_1}), SINR_{thresh}) \\ + \; & ceiling(S(i_2, j_2, k_{i_2}, k_{j_2}) - S(i_1, j_2, k_{i_1}, k_{j_2}), SINR_{thresh}) \end{aligned}$$

That defined over the entire network is:

$$SEP_{NOPC} = \sum_{i_1, i_2, j_1, j_2} SEP_{NOPC}(i_1, i_2, j_1, j_2)$$

For the separation metric with power control, it is defined as:

$$\begin{aligned} SEP_{PC}(i_1, i_2, j_1, j_2) = ceiling(\; & S(i_1, j_1, k_{i_1}, k_{j_1}) - S(i_2, j_1, k_{i_2}, k_{j_1}) \\ + \; & S(i_2, j_2, k_{i_2}, k_{j_2}) - S(i_1, j_2, k_{i_1}, k_{j_2}), 2 * SINR_{thresh}) \end{aligned}$$

Figure 7.2: Experimental map of Campus 4 (large dots: APs, small dots: clients)

Scenario	# APs	# clients	AP ant.	client ant.
Lab 2	6	6	16,35°,omni	4×35° fan beam, omni
Lab 3	13	10	all	omni
Campus 4	20	28	60°,omni	omni
City	11	9	omni	omni

Table 7.1: Experimental setup for separation metric

That defined over the entire network is:

$$SEP_{PC} = \sum_{i_1,i_2,j_1,j_2} SEP_{PC}(i_1, i_2, j_1, j_2)$$

Similarly, there are corresponding variants that consider association. For the separation without power control:

$$SEP_{NOPC,ASSOC} = \sum_{j_1 \in ASSOC(AP_{i_1}), j_2 \in ASSOC(AP_{i_2})} SEP_{NOPC}(i_1, i_2, j_1, j_2)$$

And for the separation metric with power control:

$$SEP_{PC,ASSOC} = \sum_{j_1 \in ASSOC(AP_{i_1}), j_2 \in ASSOC(AP_{i_2})} SEP_{PC}(i_1, i_2, j_1, j_2)$$

7.5 Separation Metric and Network Capacity

In Chapter 7.3, we showed that the computation of the network capacity is exponential to the number of APs in the network and two factors contribute to this complexity: the TDMA MAC

protocol that does an exhaustive search on all possible schedules and the exhaustive search on all possible antenna orientations / power levels. The separation metric, defined in the previous section, essentially uses several simplifying assumptions to significantly reduce the complexity of computing the network capacity.

In this section, we first discuss these simplifying assumptions. On one hand, these assumptions make the separation metric computationally manageable. On the other hand, they are also the necessary conditions for the separation metric to have a strong correlation with the network capacity. Based on these assumptions or conditions, we also explore when the separation metric does not work. Then we present how well the separation metric is correlated with the network capacity in three indoor environments.

7.5.1 Necessary Conditions

The first simplifying assumption is that all APs are actively transmitting at the same time. This is illustrated by the fact that the separation metric for a network is the sum of the separation metric for each element. With this assumption, we remove the exhaustive search on all possible schedules.

The second simplification is to choose antenna orientations on a pair-wise basis, i.e., on one AP and a pair of clients, on one client and a pair of APs, or on a pair of transmissions. With this simplification, we remove the exhaustive search on the antenna orientations on all senders and receivers. Using the link pair separation metric as an example, for the link pair of $AP_{i_1} \rightarrow C_{j_1}$ and $AP_{i_2} \rightarrow C_{j_2}$, the separation metric is calculated as:

$$\begin{aligned} SEP_P(i_1, i_2, j_1, j_2) = \max_{k_{i_1}, k_{i_2} \in K_{AP}, k_{j_1}, k_{j_2} \in K_C} & \quad S(i_1, j_1, k_{i_1}, k_{j_1}) - S(i_2, j_1, k_{i_2}, k_{j_1}) \\ + & \quad S(i_2, j_2, k_{i_2}, k_{j_2}) - S(i_1, j_2, k_{i_1}, k_{j_2}) \end{aligned}$$

Note that the separation metric is essentially the maximum SINR that can be achieved across antenna orientations on AP_1, AP_2, C_1, and C_2. Let $K_{AP,1}$ and $K_{C,1}$ be the antenna orientations on AP_1 and C_1 to achieve this maximum SINR. Similarly, for the link pair of $AP_{i_1} \rightarrow C_{j_1}$ and $AP_{i_3} \rightarrow C_{j_3}$, we use $K_{AP,2}$ and $K_{C,2}$ to denote the antenna orientations on AP_1 and C_1 to achieve the maximum SINR. Here, the necessary condition is that $K_{AP,1} = K_{AP_2}$ and $K_{C_1} = K_{C_2}$. That is to say, for AP_1 and C_1, there exist antenna orientations that can achieve maximum (or close-to-maximum) SINR for both $AP_{i_2} \rightarrow C_{j_2}$, and $AP_{i_3} \rightarrow C_{j_3}$.

In fact, both assumptions (or conditions) hold when the level of spatial reuse is high, e.g., for directional transmission. This indicates that the separation metric is more accurate with higher degree of spatial reuse. Due to this reason, the correlation with the separation metric and the network capacity is much lower for power control than for directional transmission, since the degree of spatial reuse that can be achieved by power control is fundamentally lower than that by directional transmission.

7.5.2 Separation Metric for Directional Antennas

Figure 7.3, Figure 7.4, and Figure 7.5 show how well the separation metric correlates with network capacity in Campus 4, Lab 3, and Lab 2 scenarios. The map for Campus 4 is shown in Figure 7.2. Note that the correlation coefficient, ρ, equaling 1 indicates a linear relationship, and $\rho = 0$ indicates no correlation (random). Each subfigure shows the scatterplot and the correlation coefficient ρ. The results suggest that (we refer to Figure 7.3 unless otherwise noted)

- Even with the simplifications, the separation metric has a positive and reasonably strong correlation with the network capacity, i.e., a higher separation metric indicates a higher capacity compared with a lower separation metric. This can be observed from the shape of the figure and the positive correlation coefficient.
- The association-aware version of the separation metric has a stronger correlation than the association-ignorant version. That is, the correlation coefficient ρ for association-aware versions are consistently higher than that for association-ignorant versions. For example, $\rho = 0.75$ for *AP,Assoc* case and $\rho = 0.64$ for *AP,NoAssoc* case. Note that even though in Lab 2, the correlation coefficient for the association-aware versions is lower than those for the association-ignorant versions, the actual shapes are in fact more linear (look for the top right part of the points).
- The AP separation metric suffers from the fact that it ignores interactions between the APs, especially when the candidate AP locations are dense, thus the correlation is relatively low. For example, $\rho = 0.64$ for *AP,NoAssoc* case is the lowest among that for all variants. In fact, any other variant does not suffer from this problem, and the correlation coefficients for all other separation metric variants are higher than that for *AP,NoAssoc* case.

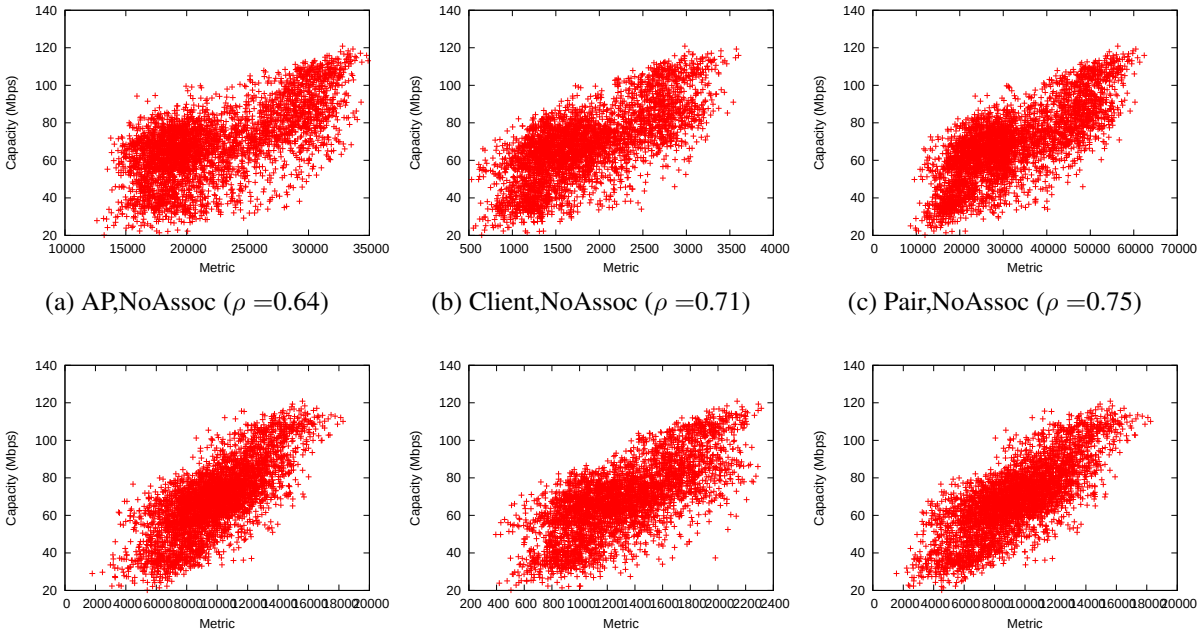
7.5.3 Shannon Capacity

Next we use Shannon theorem to explain the linear relationship between capacity and the separation metric. According to Shannon theorem:

$$C = B \log(1 + \frac{S}{N})$$

the link capacity C (the theoretical upper bound) has an almost linear relationship with the difference (in dB) of the average signal power S (in dB) through an analog communication channel to the additive white Gaussian noise of power N (in dB), i.e., the constant 1 in the formula is usually negligible.

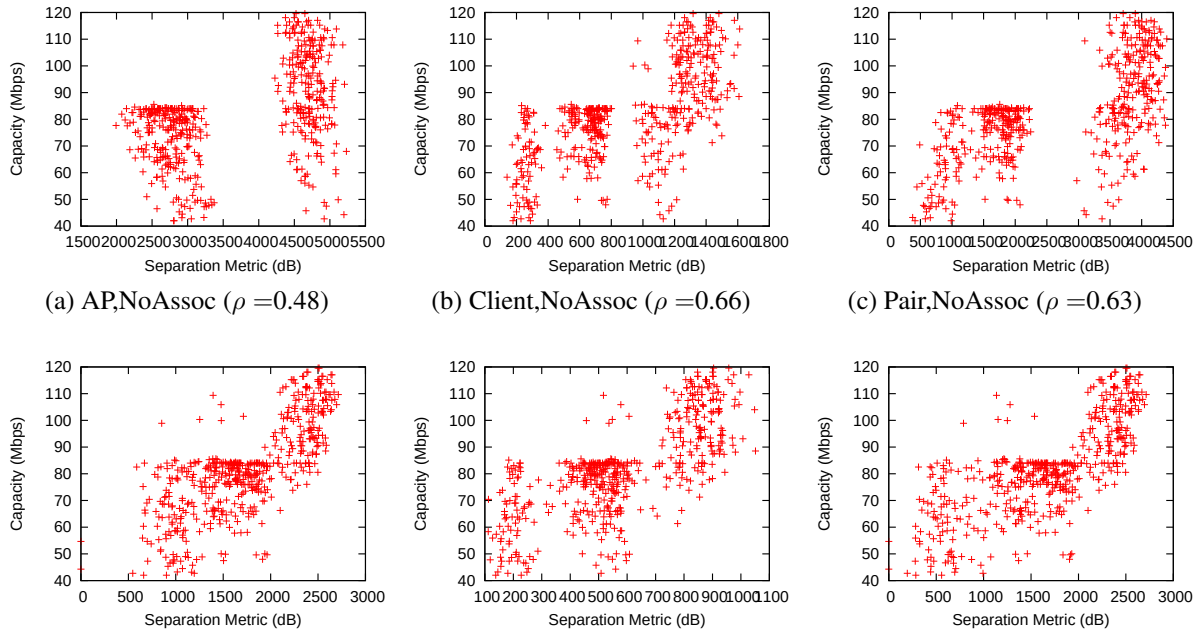
In our scenario, the separation metric basically summarizes all the SINR values in a network, thus the network capacity has strong correlation (or linear relationship) with the separation metric. Note that here we only provide an intuition, but not a proof. For example, the noise N in the Shannon theorem only refers to Gaussian noise and in the separation metric, N is replaced with interference from other sources.



(a) AP,NoAssoc (ρ =0.64) (b) Client,NoAssoc (ρ =0.71) (c) Pair,NoAssoc (ρ =0.75)

(d) AP,Assoc (ρ =0.75) (e) Client,Assoc (ρ =0.72) (f) Pair,Assoc (ρ =0.77)

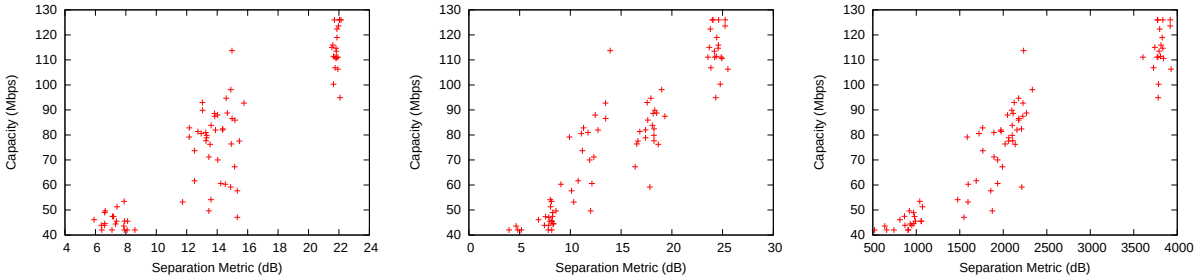
Figure 7.3: Separation metric in Campus 4 with directional APs only



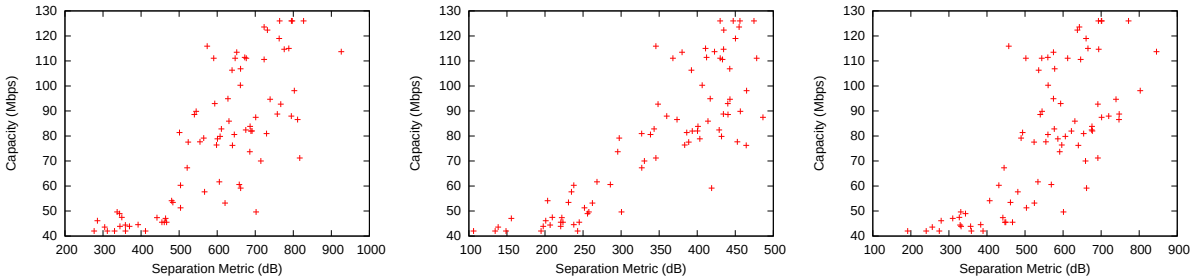
(a) AP,NoAssoc (ρ =0.48) (b) Client,NoAssoc (ρ =0.66) (c) Pair,NoAssoc (ρ =0.63)

(d) AP,Assoc (ρ =0.74) (e) Client,Assoc (ρ =0.73) (f) Pair,Assoc (ρ =0.74)

Figure 7.4: Separation metric in Lab 3 with directional APs only

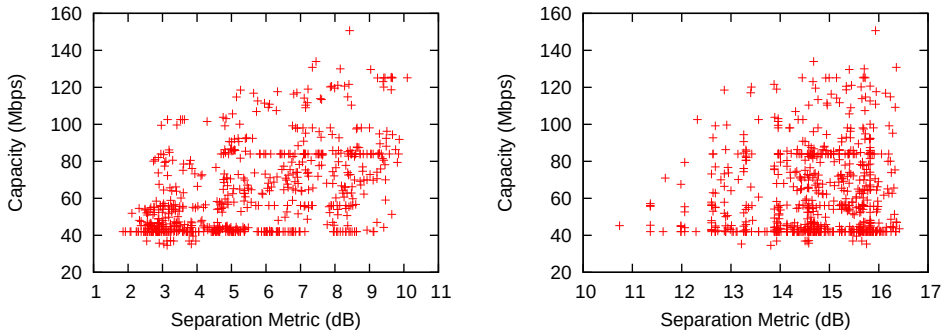


(a) AP,NoAssoc (ρ =0.91) (b) User,NoAssoc (ρ =0.92) (c) Client,NoAssoc (ρ =0.93)



(d) AP,Assoc (ρ =0.78) (e) Client,Assoc (ρ =0.85) (f) Pair,Assoc (ρ =0.73)

Figure 7.5: Separation metric in Lab 2 with directional APs and clients



(a) Pair,NoAssoc (ρ =0.49) (b) Pair,Assoc (ρ =0.64)

Figure 7.6: Separation metric in Lab 2 with power control

7.5.4 Separation Metric for Power Control

Figure 7.6 shows the correlation of the separation metric for power control and the network capacity in the City scenario. The results suggest that

- Similar to separation metric for directional antennas, the separation metric for power control is positively correlated with network capcaity.
- The correlation, however, is much weaker compared to the separation metric for directional antennas.
- The association aware version has a stronger correlation.

The reason the correlation is much lower compared with separation metric for directional antennas is due to a limitation of the separation metric. The level of spatial reuse is much lower in power control systems (usually 2 concurrent transmissions on average) compared to that in directional antenna systems (perfect isolation is some scenarios). Our separation metric is essentially a concurrent transmission metric that does not capture non-concurrent transmissions, i.e., in this case, the SINR level becomes the SNR level and is not considered in the metric. Since directional antenna systems can almost always enable concurrent transmissions, our separation metric is accurate. For power control, there are many cases with no concurrency, thus the separation metric is much less accurate.

7.5.5 Limitations

In this section, we discuss several limitations of the separation metric.

- As mentioned before, the AP metric suffers from the problem that interactions between the APs are ignored. All other variants do not have this problem.
- Also as mentioned before, the separation metric works better in scenarios with reasonable level of spatial reuse. However, even in networks with low level of spatial reuse, the correlation is still acceptable.
- A related issue with the previous limitation is that the separation metric will not work well in noise-dominated networks.
- The separation metric can only be compared in the same scenario, but not across different scenarios, since the separation metric is specific to the size of the deployment and client locations, which varies across scenarios.
- The separation metric can only be compared with the same number of APs. This is because the separation metric will increase almost linearly with more APs, but capacity will saturate at certain point and more APs will not help to improve spatial reuse. In fact, we find that the separation metric is most accurate when the capacity is almost saturated, and is less accurate otherwise.

In summary, the separation metric has a positive correlation with the network capacity if it is used to compare scenarios with the same number of APs and with the same scenario. Next, we will apply this observation to answer the questions in directional antenna and power control deployment.

Scenario	Dir. APs and clients	Dir. APs
Predicted Gain (%)	204	95
Actual Gain with 3 APs (%)	152	81
Actual Gain with 6 APs (%)	211	75

Table 7.2: Use separation metric to predict benefits of directional antennas

7.6 Using the Separation Metric

In this section, we discuss how the separation metric can help in deploying directional antenna systems. Here a site survey of signal strength measurements with directional device is necessary due to the lack of a good indoor signal propogation model. We assume a site survey has already been taken place to characterize the target indoor environment.

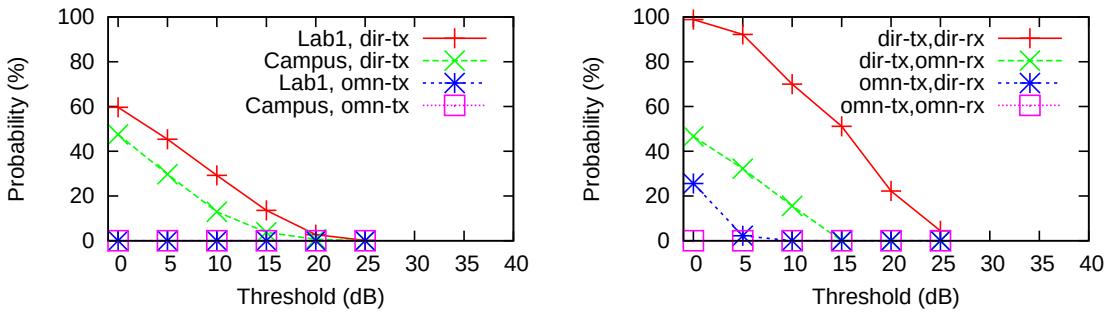
At the beginning of the chapter, we mentioned three key questions:

- Is it worth deploying directional antennas in a particular environment? Or is it worth enabling power control in a particular environment?
- What type of antenna capabilities (e.g. directional or omni-directional and beam-width) should each AP and client have?
- What are good locations for directional APs in the deployment.

7.6.1 Estimating the Benefits of Directional Transmission

The first and second questions can both be answered by estimating the performance gains from using antennas with different capabilities, i.e., the benefits of directional APs, directional clients, and directional APs and clients over omni-directional antennas. Given the key observation that the separation metric has a positive and reasonably strong correlation with network performance, we can simply look at the separation metric for different directional antenna schemes. Table 7.2 shows the predicted improvement of various directional schemes over omni-directional setup in the Lab 2 scenario, the results indicate a positive correlation between the predicted and the actual improvement. For 6 APs, the actual improvement is very close to expected. However, for 3 APs, the actual improvement is lower than expected. This is due to one of the limitations of the separation metric mentioned earlier: separation metric is most accurate when the capacity is almost saturated (at 6 APs in this scenario).

In Chapter 5, we have pointed out another benefit of directionality, flexible association, where clients in the directional antenna systems can associate with one of the multiple APs that are operating in the same channel. This allows the client to receive frames from secondary APs when the primary AP is busy. In order to look at the benefits of flexible association, we can use the client separation metric since it provides a client-centric view of multiple APs. Examining the client separation metric, we can find that for any client, the distance separation can only be positive for one AP (and negative for all other APs since it can be closest to at most one AP). On the other hand, the angular separation can be positive for more than one APs. This is exactly



(a) Campus 4 and Lab 3

(b) Lab 2

Figure 7.7: Benefiting from user association

why it make sense for clients to associate with one of the multiple APs in directional antenna systems, but it only makes sense for clients to associate with the AP with the strongest signal (within that channel) in omni-directional antenna systems.

Figure 7.7 shows for any client C_j and a pair of APs AP_{i_1} and AP_{i_2}, the probability that the client separation for both APs are higher than various thresholds, i.e., $SEP_c(i_1, i_2, j) > thresh$ and $SEP_c(i_2, i_1, j) > thresh$. The figure shows that this never happens in omni-directional antenna networks, happens a bit in directional APs or clients networks, and happens a lot for directional APs and directional clients systems. For example, it shows that there is 20% chance that a client can benefit from flexible association with a separation that is higher than 20dB for both APs, indicating that the client can support at least 36Mbps even it associates with the secondary AP and the primary AP is busy transmitting to other receivers (both transmitting simultaneously in the same channel).

7.6.2 Estimating the Benefits of Power Control

To estimate the benefits of power control, we can also compare the separation metric. In the Lab 2 scenario, the improvement in separation metric is 13% from the association ignorant version, and is 21% from the association aware version. The actual capacity improvement is 31%. Note that the prediction is much less accurate than that in the directional antenna case due to the fact that the separation metric is less accurate for power control, as described in Chapter 7.5.

7.6.3 Finding the Desirable Locations for Directional APs

Suppose the separation metric indicates that deploying directional APs is worthwhile in a particular environment, there is still the challenge of identifying desirable locations for directional APs. Here, we present how to use the separation metric to make this decision. To simplify presentation, we only focus on performance in terms of spatial reuse of spectrum, but ignore other practical factors such as ease of deployment. The problem is to choose the m AP locations out of M candidate locations that maximize spatial reuse. Note that one limitation of the separation metric is that it is unable to find the best m automatically. The number m must be determined by

the network operator according to the expected amount of traffic. The brute force approach using SNR measurements is to run simulation for all possible $choose(m, M)$ AP placement strategies. Thus, the whole process incurs a cost of $O(choose(m, M) * K_{ap}^{m} K_{c}^{m})$, which is simply too slow, especially for a large deployment.

Using the separation metric, we can simply choose the set of APs that maximizes the AP separation metric, and note that this is an incremental and greedy process. In case the predication of the separation metric turns out to be bad, we can also use the simulation to verify the choice of AP placement, by simulating the AP placement with the highest separation metric, the next highest, etc., until the k-th highest. We call this approach the MaxSep approach.

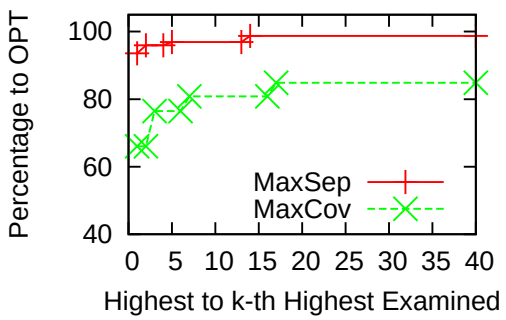
For comparison, we look at the AP placement strategy to maximizing coverage, we call this approach the MaxCov approach, which chooses the set of APs that are maximizes the minimum signal strength from all clients. Similarly, simulation can be used here to verify the AP placement with the highest distance, the next highest, etc., until the k-th highest.

Figure 7.8 shows the effectiveness of both the MaxSep and MaxCov placement strategies. Each figure shows the best capacity that can be achieved after examining k highest scenarios as a percentage of the optimal capacity. For example, Figure 7.8(a) shows that for the Campus scenario, MaxCov can only achieve 67% of optimal performance after first trial, and MaxSep can in fact achieve 94% of optimal performance. The results show that the MaxCov approach does not work well in the Campus scenario and Lab2 scenario, and MaxSep approach works reasonably well in all scenarios.

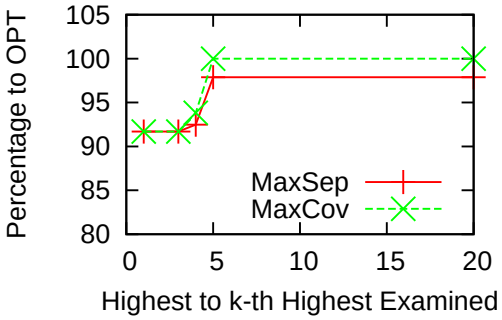
Recall that the first limitation of the AP metric is that it only works well in sparse AP candidate locations, e.g., in our scenarios APs are at least 2 meters from each other. In order to illustrate this, we also carried out an experiment in the Campus1 scenario with very dense candidate AP locations (60 AP locations with about 20-60cm distance between two APs). Figure 7.8 shows the performance of planning algorithm that uses the AP metric and the client metric. The results show that the client metric works much better in this scenario. This indicates that the AP metric should be used when the candidate AP locations are sparse, e.g., to upgrade an existing omni-directional deployment. However, when the candidate AP locations are dense, client metric is in fact a better choice.

7.6.4 Stability

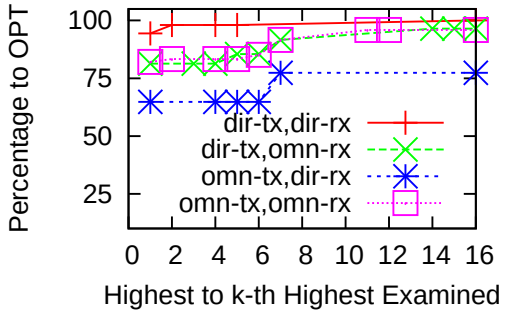
For planning purposes, we also look at how will small perturbations in time and distance change the separation metric and the network capacity. Figure 7.9(a)&(b) show how the AP metric change over small perturbations in AP location and time. The results suggest that small perturbations affect separation metric only slightly. Figure 7.9(c) shows the network capacity before and after we move all the clients in the Campus scenario by 1-2 meters. The results suggest that even though all the users move a bit, the capacity does not change significantly. This is consistent with the observation that small movements result in small changes observed in the separation metric.



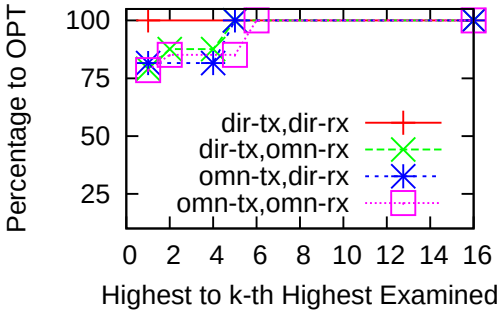
(a) Campus



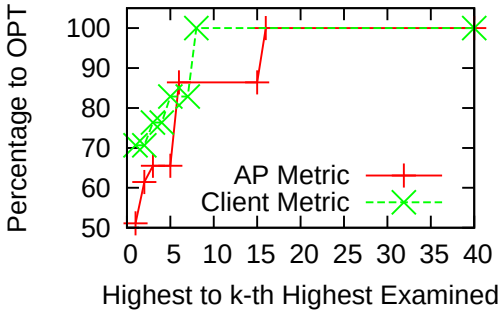
(b) Lab1



(c) Lab2 (MaxCov)

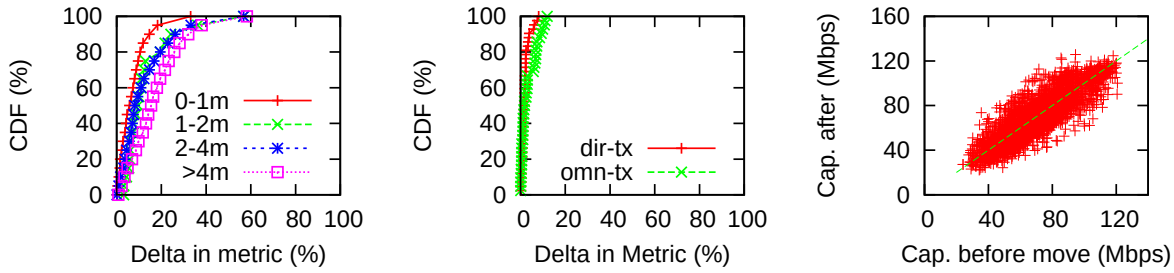


(d) Lab2 (MaxSep)



(e) Dense APs (MaxSep with AP / client metrics)

Figure 7.8: Effectiveness of the planning algorithms



(a) SEP_{ap},loc (b) SEP_{ap},time (c) capacity,loc

Figure 7.9: Changes in separation metric and network capacity with small perturbations in time and space

7.7 Summary and Discussion

In the first part of this chapter, we presented three major factors that determine the effectiveness of directional transmission and power control in any particular environment. In the second part of this chapter, we proposed the separation metric, which is a simple measurement based metric for an environment that is strongly correlated with the network capacity of directional antenna systems in that environment. The proposed separation metric has limitations such as only enabling comparison between deployment configurations with the same number of APs. Despite its limitations, we show that the separation metric is very useful in addressing key problems in planning wireless networks: 1) predicting the benefits of directional antennas in indoors environments, and 2) guiding the placement of directional APs.

Chapter 8

Conclusions and Future Work

8.1 Contributions

In this dissertation, we showed the validity of the following thesis:

> *There exist practical and lightweight designs for both directional transmission and power control in which nodes coordinate with each other to effectively optimize network-wide spatial reuse in chaotic and dense wireless networks even in rich scattered indoor environments.*

In addition to showing the validity of the thesis, we also made the following contributions in this dissertation:

- We show that both directional transmission and power control can be very effective, i.e., significantly improve spatial reuse over naive solutions, even in rich scattered indoor environments if the systems are carefully designed.
- Using the SINR model, we greatly reduce the number of measurements to choose the right antenna orientations and power levels in our systems to optimize spatial reuse in a quick, lightweight, and practical fashion.
- We design and build a timeslot and timeslot reservation based MAC protocols (both centralized and decentralized) to exploit the spatial reuse opportunities, i.e., to prevent (directional) hidden terminal and (directional) exposed terminal problems, based on the observations that CSMA based protocols interact poorly with both directional transmission and power control.
- Using a simple SINR based metric, the separation metric (including distance and angular separation), we can quickly evaluate the effectiveness of directional transmission and power control in any particular environment, and can guide the placement of directional APs in that environment.

Next we summarize key insights in each contribution.

8.1.1 Effectiveness of Directional Transmission and Power Control

Based on measurements in several indoor scenarios, we show that both directional transmission and power control can be useful even rich scattered indoor environments.

The key insight is that the effectiveness of directional transmission and power control can only be exploited on two conditions: 1) when wireless nodes make coordinated decisions to choose antenna orientations or power levels, and 2) has a good non-CSMA MAC protocol. As far as we know, we are the first to provide this insight.

8.1.2 Using the SINR model

The naive coordinated solution to choose the antenna orientations or power levels are simply too slow in practice. We propose to use the SINR model to reduce the number of measurements.

The key insight is two-fold: 1) the use of SINR model can significant reduce the number of measurements to choose the right antenna orientations or power levels, and 2) how to modify the SINR model to make it more practical, e.g., incorporating multiple data rates and intermediate loss rates and how to automatically calibrate each card.

8.1.3 Timeslot MAC protocol

The key insight is that CSMA based MAC protocols interact poorly with both directional transmission and power control. Another insight is that there exists timeslot based MAC protocol (both centralized and distributed) that can both exploit the capabilities of directional transmission and power control, and can co-exist with other CSMA based solutions.

8.1.4 Directional Antenna Systems

Aside from the effectiveness of directional antenna systems and the MAC protocol, there are three more insights in directional antenna systems. First, we find that the configuration of directional APs and directional clients is fudamentally more powerful than the configuration of directional APs only, especially when the clients are clustered as in conference rooms. Second, we find that there exists very practical antenna setup on wireless clients that can exploit the benefits of directional antennas. Third, we find that flexible association can be implemented in directional antenna systems to balance AP load and to improve the network capacity.

8.1.5 Power Control Systems

In addition to showing the effectiveness of power control systems and the MAC protocol, there are two more insights in power control systems. First, we find that power levels need to be tuned on a per-link basis to maximize spatial reuse. Unfortunately, this problem is conditional NP-hard. Second, we show that by choosing the middle power level among the range of power levels, a greedy algorithm works reasonably well in practice. This is due to the fact that choosing the middle power level leaves some slack in accommodating interference level from future

transmissions and some slack for the future transmissions to avoid being interfered by the current transmission.

8.2 Limitations and Future Work

In this section, we present several limitations of our systems and an agenda for future work.

8.2.1 Expectations for Future Wireless Cards

We first summarize several implementation issues due to the limitations of existing wireless cards and firmware, and what features are needed for future wireless cards and firmware to effectively run our systems.

- Antenna connectors: Currently most cards have only 2 antenna connectors. In Speed, each Speed client needs one wireless card with (at least) 5 antenna connectors. Also, currently we cannot obtain the RSSI readings on both antenna connectors for a single incoming frame. However, since the card can automatically choose the antenna connector with the stronger signal strength, indicating that this information is available. It is desirable that this information be exposed to firmware / HAL.
- Remove frames from hardware queues: The current API between the driver and the firmware / HAL is that the driver sends the frames to the hardware queue, but there is no API for removing a frame from the queues. It is desirable that frames can be removed, which is necessary to implement the MAC protocol in Speed and Opera.

8.2.2 Characterizing Benefits in UHF and 60GHz

As discussed in Chapter 1.3.1, our conclusions are only validated in 2.4GHz bands, and we believe that they can be applied to UHF and 60GHz as well. However, in order to validate this, we would like to quantify the benefits of directional transmission and power control in those bands.

8.2.3 More Practical Directional Antenna Setup

In Chapter 5, we show that the antenna setup with a phased array antenna on the APs and a small number of patch antennas on wireless clients, works well in infrastructure wireless networks, e.g., enterprise networks and wireless hot spots. This particular antenna setup, however, may not be appropriate for device-to-device application scenarios such as those enabled by Wi-Fi Direct [7], because of two reasons:

- Existing commercial phased array antennas are quite expensive, making them unsuitable on every day consumer devices.
- A more fundamental limitation is the size of the phased array antennas. They are too large (fundamentally) for the small devices.

In Chapter 5, we observe that a small number of patch antennas suffice on wireless clients due to rich scattering in indoor environments. Since this observation is made with the omni-directional antennas on the APs, so further measurements is necessary to characterize the scenario when both APs and clients use a small number of patch antennas. Ideally, to obtain the complete set of measurements, the measurement study will involve two (or more) turntables that mount various different antennas with different beamwidth. The notion of directionality cost (introduced in Chapter 5) will be useful here to characterize the degree of directionality.

Another related future work is antenna engineering, while in Speed, we use commercial off-the-shelf patch antennas on wireless clients, it is possible to design specific type of patch antennas on these clients. Several parameters that require further study are

- The size of the antenna vs. the beamwidth: Generally, the beamwidth will be smaller if the size of the antenna increases, and for different devices (especially different shapes), the ideal size or beamwidth may be different. This is also determined by the frequency bands used, i.e., lower frequency bands means larger antennas for the same beamwidth.
- The beamwidth in horizontal vs. vertical planes: Generally, with the size of the antenna relatively fixed, we can still change the shape of the antenna to change the beamwidth in one plane versus the other plane. In many scenarios, the performance of a fan beam antenna, i.e., that with a narrower beam on horizontal plane and a wider beam on the vertical plane, is better (or can provide a higher level of isolation) than that of the original antenna, i.e., that have the same beams on both planes.
- The ability to enable multiple sectors: Due to the limitation of our hardware setup, we can only use one patch antenna at a time. But in some sectorized antennas, multiple sectors can be enabled at the same time [113], which effectively produces more antenna patterns (with different beamwidth). Such capability may become handy in several scenarios, e.g., handling dynamics as described in the next section.

8.2.4 Dynamics and Beamwidth

All our systems are designed to optimize spatial reuse, with the consideration of nomadic usages. In evaluating the systems, we also show that our systems can handle limited level of dynamics, e.g., node mobility with walking speed. However, we did not evaluate our systems under high level of dynamics such as lots of people moving around.

In fact, beamwidth adaptation may be used to allow our systems to better handle dynamics or even handle high level of dynamics. In beamwidth adaptation, wireless nodes monitor the level of dynamics and adapt the beamwidth to the environment: if the environment is highly dynamic, use wide beams, if the environment is stable, use narrow beams, and if the environment is extremely dynamic, use omni-directional patterns. The key is how to determine the most appropriate beams for any environment. Using narrow beam antennas in highly dynamic environments will lead to bad performance when the environment has changed but the antenna configuration has not (causing the systems to operate in sub-optimal states). By making measurements more aggressively, it may incur very high cost of measurement traffic. Using wide beam antennas will reduce the number of configuration changes and the number of measurements. However, it will also reduce the level of spatial reuse that can be achieved in that particular environment. Thus

finding the most appropriate beamwidth is an interesting and challenging task.

Another potential direction to better handle dynamics is to predict the behavior of the nodes based on history. One possible feature is to adapt the measurement interval or the set of directions to measure based on history, e.g., only scan the directions that are likely to change. Also, even in rich scattered indoor environments, signal strength may not change abruptly during node movement or change of orientation [12]. This indicates that we may be able to predict the signal strength patterns, e.g., Midas is a system that predicts the signal strength at various receive sectors of a directional receiver based on history.

8.2.5 MIMO Beamforming and Pattern Adaptation

While commercial phased array antennas are expensive, MIMO technologies are getting less expensive and gaining more popularity. This is true for both 2.4GHz and 60GHz. The idea of our systems, i.e., nodes coordinate with each other to maximize spatial reuse, can also be applied to MIMO antenna settings. In fact, MIMO radios can obtain much more information on the multipath channel condition between the sender and the receiver, with even lower overhead. There are several implications of applying our idea to MIMO beamforming.

First, using MIMO radios, the measurement overhead can be reduced to minimum. In MIMO networks, only one frame (or several frames) is needed to characterize the multipath channel between an AP to all the clients, incurring a cost of $O(M)$. Note that in our systems, the number of measurements is $O(M * K_{AP})$.

Second, given the channel path condition information, i.e., channel matrices, an optimization can be applied to generate some particular antenna pattern that can maximize signal on the intended receiver and can have nulls towards the unintended receivers. Techniques such as singular value decomposition (SVD) will be useful.

Third, aside from beamforming, spatial multiplexing is also another important usage of MIMO radios. Spatial multiplexing can send multiple (k') data streaming to maximize link capacity with k antennas. While in theory, $k' = k$, but in practice, k' will be increased to a certain point where further increasing k will not increase k'. In that case, it is desirable to allocate some antennas for spatial multiplexing and others for beamforming to increase spatial reuse. However, performing beamforming will also affect k', i.e., it may reduce the number of paths between the sender and the receiver. The question is in any particular environment, what is the most appropriate number of antennas for spatial multiplexing and what is the most appropriate number for beamforming.

The authors in [68] propose a very interesting technique to estimate the multipath channel between a phased array antenna and an off-the-shelf omni-directional client. Using this estimation, the directional APs can form adaptive antenna patterns that reinforce the signal strength at the receiver. With such more detailed information of the channel condition, our systems can be extended to use non-default antenna patterns, thus further increasing spatial reuse.

8.2.6 Handling Bursty Traffic Patterns

In Chapter 6, we presented one limitation of our current implementation of the timeslot based MAC protocol that it does not handle bursty traffic very well. This is because in our current

implementation, one AP reserved timeslot can only be used for the AP to transmit frames to one particular client. Thus if the traffic from the clients are bursty, the reserved timeslot may become largely unused. In Chapter 6, we discussed several future steps to better utilize the timeslots by avoid wasting the reserved timeslots, including having APs aggregate traffic for the client, or allowing the APs to cancel timeslot reservations.

8.2.7 Non-infrastructure Networks

Our systems primarily target single-hop infrastructure wireless networks not only because infrastructure networks are the most prevalent type of networks, but also because the problem in the one-hop networks is already difficult enough. That said, our ideas might be applied to multi-hop networks such as mesh or ad-hoc networks, with much greater complexity.

Basically, optimization metric of network capacity does not amount to the sum of link throughputs; instead, it is the sum of path throughputs. Path throughput depends on the routing protocol running, thus joint optimization of routing and antenna orientations and/or power levels is necessary. This significantly complicates the problem.

8.2.8 Sectorized APs

While we use phased array antennas on APs, sectorized directional antennas can also be used on APs. Both techniques have pros and cons:

- Phased array antennas can be oriented to any orientation and can form irregular antenna patterns. Sectorized antennas usually have a fixed number of fixed shaped antenna patterns.
- Usually there is only one radio on the phased array antennas, while there are multiple radios on the sectorized antennas.

One possible future work is to quantify these pros and cons through a more comprehensive study of these two techniques.

Bibliography

[1] *IEEE 802.11n standard*. 4.5.5

[2] *Atheros (www.atheros.com)*. 2.2.2, 3.2.1

[3] *Fidelity Comtech (www.fidelity-comtech.com)*. 1.1.1, 2.2.1, 2.2.3, 4.4.1, 4.5.3, 6

[4] *GNURadio (www.gnuradio.org)*. 3.3

[5] *Ruckus Wireless (www.ruckuswireless.com)*. 1.1.1, 2.2.3, 4.5.3, 6

[6] *SiBEAM: Wireless Beyond Boundaries (www.sibeam.com)*. 2.4.4, 7.2.3

[7] *Wi-Fi Alliance (www.wi-fi.org)*. 8.2.3

[8] *Xirrus High Performance Wi-Fi (www.xirrus.com)*. 1.1.1, 2.2.3, 4.5.3, 6

[9] Daniel Aguayo, John Bicket, Sanjit Biswas, Glenn Judd, and Robert Morris. Link-level measurements from an 802.11b mesh network. *SIGCOMM Comput. Commun. Rev.*, 34 (4):121–132, 2004. ISSN 0146-4833. 3.2.2

[10] Nabeel Ahmed, Usman Ismail, Srinivasan Keshav, and Konstantina Papagiannaki. Online estimation of RF interference. In *CoNEXT*, 2008. 3.1, 3.3

[11] Aditya Akella, Glenn Judd, Srinivasan Seshan, and Peter Steenkiste. Self management in chaotic wireless deployments. In *MobiCom*, 2005. 1, 3.3, 6, 6, 6.1.1

[12] Ardalan Amiri Sani, Lin Zhong, and Ashutosh Sabharwal. Directional antenna diversity for mobile devices: characterizations and solutions. In *MobiCom*, 2010. 1.1.1, 1.1.3, 4.5, 4.5.3, 8.2.4

[13] Eric Anderson, Gary Yee, Caleb Phillips, Douglas Sicker, and Dirk Grunwald. Commodity ar52xx-based wireless adapters as a research platform. Technical report, University of Colorado at Boulder, Department of Computer Science, April 2008. 4.3.2

[14] Erdal Arikan. Some complexity results about packet radio networks. *IEEE Transactions on Information Theory*, 30:681–685, 1984. 4.2.4

[15] Ehsan Aryafar, Narendra Anand, Theodoros Salonidis, and Edward W. Knightly. Design and experimental evaluation of multi-user beamforming in wireless lans. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, MobiCom '10, pages 197–208, 2010. 4.5.5

[16] Paramvir Bahl, Ranveer Chandra, and John Dunagan. Ssch: slotted seeded channel hopping for capacity improvement in ieee 802.11 ad-hoc wireless networks. In *Proceedings of the 10th annual international conference on Mobile computing and networking*, Mobi-

Com '04, pages 216–230, 2004. 5.2.2

[17] Balabhaskar Balasundaram and Sergiy Butenko. Graph domination, coloring and cliques in telecommunications. 4.5.4

[18] Lichun Bao and J. J. Garcia-Luna-Aceves. Receiver-oriented multiple access in ad hoc networks with directional antennas. *Wirel. Netw.*, 11(1-2), 2005. 4.5.1

[19] Yigal Bejerano, Seung-Jae Han, and Li (Erran) Li. Fairness and load balancing in wireless lans using association control. In *MobiCom '04: Proceedings of the 10th annual international conference on Mobile computing and networking*, pages 315–329, New York, NY, USA, 2004. ACM. ISBN 1-58113-868-7. doi: http://doi.acm.org/10.1145/1023720.1023751. 5.2.2

[20] Marc Blanco, Ravi Kokku, Kishore Ramachandran, Sampath Rangarajan, and Karthik Sundaresan. On the effectiveness of switched beam antennas in indoor environments. In *PAM*, 2008. 2.2.1, 2.6, 7, 7.1.3

[21] Holger Boche, Marcin Wiczanowski, and Slawomir Stanczak. Unifying view on min-max fairness, max-min fairness, and utility optimization incellular networks. *EURASIP J. Wirel. Commun. Netw.*, 2007:5–5, January 2007. 1.1.3, 2.1, 4.5.4

[22] Gurashish Brar, Douglas M. Blough, and Paolo Santi. Computationally efficient scheduling with the physical interference model for throughput improvement in wireless mesh networks. In *MobiCom*, 2006. 3.1, 3.3, 3.3, 4.2.4, 4.2.4, 4.5.4

[23] Micah Z. Brodsky and Robert T. Morris. In defense of wireless carrier sense. In *Proceedings of the ACM SIGCOMM 2009 conference on Data communication*, SIGCOMM '09, 2009. 2.4.3

[24] Ioannis Broustis, Jakob Eriksson, Srikanth V. Krishnamurthy, and Michalis Faloutsos. Implications of power control in wireless networks: A quantitative study. In *PAM*, 2007. 2.6

[25] Ioannis Broustis, Konstantina Papagiannaki, Srikanth V. Krishnamurthy, Michalis Faloutsos, and Vivek Mhatre. Mdg: measurement-driven guidelines for 802.11 wlan design. In *MobiCom*, 2007. 5.2.2

[26] Michael Buettner, Eric Anderson, Gary Yee, Dola Saha, Anmol Sheth, Douglas Sicker, and Dirk Grunwald. A phased array antenna testbed for evaluating directionality in wireless networks. In *MobiEval*, 2007. 2.2.1

[27] J.M. Carey and D. Grunwald. Enhancing wlan security with smart antennas: a physical layer response for information assurance. *VTC*, pages 318–320 Vol. 1, 2004. 2.2.1

[28] V. Chandrasekhar and J.G. Andrews. Uplink capacity and interference avoidance for two-tier cellular networks. In *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, pages 3322 –3326, 2007. 6.1.3

[29] V. Chandrasekhar, J. Andrews, and A. Gatherer. Femtocell networks: a survey. *Communications Magazine, IEEE*, 46(9):59 –67, 2008. 6.1.3

[30] Vikram Chandrasekhar, Jeffrey G. Andrews, Tarik Muharemovic, Zukang Shen, and Alan Gatherer. Power control in two-tier femtocell networks. *Trans. Wireless. Comm.*, 8:4316–

4328, August 2009. 6.1.3

[31] Kameswari Chebrolu and Bhaskaran Raman. FRACTEL: a fresh perspective on (rural) mesh networks. In *NSDR*, 2007. 4.5.1

[32] Young-June Choi and K.G. Shin. Power-adjusted random access to a wireless channel. *INFOCOM*, 2008. 6, 6.1.1

[33] Romit Roy Choudhury and Nitin H. Vaidya. Deafness: A MAC problem in ad hoc networks when using directional antennas. *ICNP*, 2004. 4.5.1

[34] Romit Roy Choudhury, Xue Yang, Ram Ramanathan, and Nitin H. Vaidya. Using directional antennas for medium access control in ad hoc networks. In *MobiCom*, 2002. 4.5.1

[35] T Cover and J Thomas. *Elements of Information Theory*. J Wiley and Sons Inc, 1991. 4.5.5

[36] Saumitra M. Das, Dimitrios Koutsonikolas, Y. Charlie Hu, and Dimitrios Peroulis. Characterizing multi-way interference in wireless mesh networks. In *ACM MobiCom International Workshop WiNTECH*, 2006. 3.1.1

[37] A. Demers, S. Keshav, and S. Shenker. Analysis and simulation of a fair queueing algorithm. *SIGCOMM Comput. Commun. Rev.*, 19:1–12, August 1989. 2.1

[38] M. Dinitz. Distributed algorithms for approximating wireless network capacity. In *INFOCOM, 2010 Proceedings IEEE*, pages 1 –9, 2010. 1.1.3, 6.6

[39] Petar Djukic and Shahrokh Valaee. Link scheduling for minimum delay in spatial re-use tdma. In *IN PROCEEDINGS OF INFOCOM*, pages 28–36, 2007. 4.5.4

[40] Mesut Ali Ergin, Kishore Ramachandran, and Marco Gruteser. Understanding the effect of access point density on wireless lan performance. In *MobiCom*, 2007. 6

[41] H. Fattah and C. Leung. An overview of scheduling algorithms in wireless multimedia networks. *Wireless Communications*, 9, October 2002. 4.5.4

[42] Jason A. Fuemmeler, Nitin H. Vaidya, and Venugopal V. Veeravalli. Selecting transmit powers and carrier sense thresholds in csma protocols for wireless ad hoc networks. In *WICON*, 2006. 3.1, 3.3, 6

[43] Sachin Ganu, Haris Kremo, Richard Howard, and Ivan Seskar. Addressing repeatability in wireless experiments using orbit testbed. In *TRIDENTCOM*, 2005. 3.2.2

[44] D. Gesbert, M. Kountouris, R.W. Heath, Chan-Byoung Chae, and T. Salzer. Shifting the MIMO paradigm. *Signal Processing Magazine, IEEE*, 2007. 4.5.5

[45] Domenico Giustiniano, David Malone, Douglas J. Leith, and Konstantina Papagiannaki. Experimental assessment of 802.11 mac layer channel estimators. In *IEEE Communications Letters*, 2007. 5.2.2

[46] Shyamnath Gollakota and Dina Katabi. Zigzag decoding: combating hidden terminals in wireless networks. *SIGCOMM Comput. Commun. Rev.*, 38(4):159–170, 2008. ISSN 0146-4833. doi: http://doi.acm.org/10.1145/1402946.1402977. 3.3

[47] Shyamnath Gollakota, Samuel David Perli, and Dina Katabi. Interference alignment and

cancellation. *SIGCOMM Comput. Commun. Rev.*, 39:159–170, August 2009. 3.3

[48] Ashutosh Deepak Gore, Srikanth Jagabathula, and Abhay Karandikar. On high spatial reuse link scheduling in stdma wireless ad hoc networks. In *IN IEEE GLOBECOM*, 2007. 4.5.4

[49] Piyush Gupta and P. R. Kumar. The Capacity of Wireless Networks. In *IEEE Transactions on Information Theory, vol. IT-46*, pages 388–404, 2000. 3.1, 3.3

[50] Daniel Halperin, Thomas Anderson, and David Wetherall. Taking the sting out of carrier sense: interference cancellation for wireless lans. In *Proceedings of the 14th ACM international conference on Mobile computing and networking*, MobiCom '08, pages 339–350, 2008. 3.3

[51] Martin Heusse, Franck Rousseau, Romaric Guillier, and Andrzej Duda. Idle sense: an optimal access method for high throughput and fairness in rate diverse wireless lans. *SIGCOMM Comput. Commun. Rev.*, 35:121–132, August 2005. 2.1

[52] Wen hwa Liao, Yu-Chee Tseng, and Kuei-Ping Shih. A tdma-based bandwidth reservation protocol for qos routing in a wireless mobile ad hoc network. *Communications, ICC 2002. IEEE International Conference on*, 5:2002, 2002. 4.5.4

[53] Kamal Jain, Jitendra Padhye, Venkata N. Padmanabhan, and Lili Qiu. Impact of interference on multi-hop wireless network performance. In *MobiCom*, 2003. 3.1, 3.3, 4.5.4

[54] R. Jain, D. Chiu, and W. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems. Technical Report TR-301, DEC Research, September 1984. 2.1

[55] Kyle Jamieson, Bret Hull, Allen Miu, and Hari Balakrishnan. Understanding the real-world performance of carrier sense. In *E-WIND*, 2005. 5.3.2, 5.3.5

[56] Lujun Jia, Xin Liu, G. Noubir, and R. Rajaraman. Transmission power control for ad hoc wireless networks: throughput, energy and fairness. *WCNC*, 2005. 6, 6.1.1

[57] Han-Shin Jo, Cheol Mun, June Moon, and Jong-Gwan Yook. Interference mitigation using uplink power control for two-tier femtocell networks. *Wireless Communications, IEEE Transactions on*, 8(10):4906 –4910, 2009. 6.1.3

[58] Rick A. Jones. netperf: A network performance benchmark, 1993. 3.2

[59] Glenn Judd and Peter Steenkiste. Using emulation to understand and improve wireless networks and applications. In *NSDI*, 2005. 3.2, 3.2.2

[60] E. Jung and N. Vaidya. A power control mac protocol for ad-hoc networks. *MOBICOM*, 2002. 6, 6.1.1

[61] Dina Katabi, Mark Handley, and Charlie Rohrs. Congestion control for high bandwidth-delay product networks. *SIGCOMM Comput. Commun. Rev.*, 32:89–102, August 2002. 2.1

[62] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM J. Comput.*, 37(1), 2007. 6.6

[63] Kyu-Han Kim, Alexander W. Min, and Kang G. Shin. Sybot: an adaptive and mobile spectrum survey system for wifi networks. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, MobiCom '10, pages 293–304, 2010. 7.3

[64] Tae-Suk Kim, Jennifer C. Hou, and Hyuk Lim. Improving spatial reuse through tuning transmit power, carrier sense threshold, and data rate in multihop wireless networks. In *MobiCom*, 2006. 3.1, 3.3, 6, 6, 6.1.1, 6.1.2, 6.2.1

[65] Young-Bae Ko, Vinaychandra Shankarkumar, and Nitin H. Vaidya. Medium access control protocols using directional antennas in adhoc networks. In *INFOCOM*, 2000. 1.1.1, 1.1.2, 4.5.1, 4.5

[66] Thanasis Korakis, Gentian Jakllari, and Leandros Tassiulas. A MAC protocol for full exploitation of directional antennas in ad-hoc wireless networks. In *MobiHoc*, 2003. 4.5.1

[67] Ulas C. Kozat, Iordanis Koutsopoulos, and Leandros Tassiulas. A framework for cross-layer design of energy-efficient communication with qos provisioning in multi-hop wireless networks. 2004. 4.5.4

[68] Sriram Lakshmanan, Karthikeyan Sundaresan, Sampath Rangarajan, and Raghupathy Sivakumar. Practical beamforming based on rssi measurements using off-the-shelf wireless clients. In *IMC '09: Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pages 410–416, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-771-4. doi: http://doi.acm.org/10.1145/1644893.1644942. 2.2.1, 4.5.2, 8.2.5

[69] Sriram Lakshmanan, Karthik Sundaresan, Sampath Rangarajan, and Raghupathy Sivakumar. The myth of spatial reuse with directional antennas in indoor wireless networks. In *PAM*, 2010. 7, 7.1.1

[70] J. Lee, W. Kim, S.-J. Lee, D. Jo, J. Ryu, T. Kwon, and Y. Choi. An experimental study on the capture effect in 802.11a networks. In *ACM WiNTECH*, 2007. 3.2, 3.2.1, 3.3

[71] Tong Li, Dan Baumberger, and Scott Hahn. Efficient and scalable multiprocessor fair scheduling using distributed weighted round-robin. *SIGPLAN Not.*, 44:65–74, February 2009. 2.1

[72] Xi Liu, Srinivasan Seshan, and Peter Steenkiste. Understanding directional antennas in dense indoor deployments. . 7

[73] Xi Liu, Anmol Sheth, Michael Kaminsky, Konstantina Papagiannaki, Srinivasan Seshan, and Peter Steenkiste. Opera: Optimizing spatial reuse through fine-grained power control. . 1, 1.1, 2.2.3, 6

[74] Xi Liu, Srini Seshan, and Peter Steenkiste. Interference-Aware Transmission Power Control for Dense Wireless Networks. In *Proceedings of the Annual Conference of ITA*, 2007. **NOTE:** This was an unrefereed, internal venue for the ITA project. It does not preclude later publications that overlap in content. 3.3, 6.3.2

[75] Xi Liu, Anmol Sheth, Michael Kaminsky, Konstantina Papagiannaki, Srinivasan Seshan, and Peter Steenkiste. Dirc: increasing indoor wireless capacity using directional antennas. *SIGCOMM Comput. Commun. Rev.*, 39(4):171–182, 2009. 1, 1.1, 2.2.3, 4

[76] Xi Liu, Anmol Sheth, Michael Kaminsky, Konstantina Papagiannaki, Srinivasan Seshan, and Peter Steenkiste. Pushing the envelope of indoor wireless spatial reuse using directional access points and clients. In *MobiCom*, 2010. 1, 1.1, 2.2.3, 5

[77] Xin Liu, Edwin K. P. Chong, and Ness B. Shroff. Opportunistic transmission scheduling with resource-sharing constraints in wireless networks. *IEEE Journal on Selected Areas in Communications*, 19:2053–2064, 2001. 4.5.4

[78] Haiyun Luo, Songwu Lu, Vaduvur Bharghavan, Jerry Cheng, and Gary Zhong. A packet scheduling approach to qos support in multihop wireless networks. In *Mob. Netw. Appl*, pages 193–206, 2002. 4.5.4

[79] Ratul Mahajan, Maya Rodrig, David Wetherall, and John Zahorjan. Analyzing the MAC-level behavior of wireless networks in the wild. In *SIGCOMM*, 2006. 1

[80] Justin Manweiler, Naveen Santhapuri, Souvik Sen, Romit Roy Choudhury, Srihari Nelakuditi, and Kamesh Munagala. Order matters: transmission reordering in wireless networks. In *MobiCom*, 2009. 3.2.1, 3.3

[81] V. Mhatre, K. Papagiannaki, and F. Baccelli. Interference mitigation through power control in high density 802.11 wlans. In *Infocom*, 2007. 1.1.2, 1.1.3, 3.1, 3.3, 6, 6, 6.1.1, 6.1.2, 6.2.1, 4, 6.3.2

[82] J.P. Monks, V. Bharghavan, and W.-M.W. Hwu. A power controlled multiple access protocol for wireless packet networks. *INFOCOM*, 2001. 6, 6.1.1

[83] A. Muqattash and M. Krunz. Power controlled dual channel (pcdc) medium access protocol for wireless ad hoc networks. *INFOCOM*, 2003. 6, 6.1.1

[84] Alaa Muqattash and Marwan Krunz. A single-channel solution for transmission power control in wireless ad hoc networks. In *MobiHoc*, 2004. 6, 6.1.1

[85] Swetha Narayanaswamy, Vikas Kawadia, R. S. Sreenivas, and P. R. Kumar. Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the compow protocol. In *in European Wireless Conference*, 2002. 6, 6.1.1, 6.1.2

[86] V. Navda, R. Kokku, S. Ganguly, and S. Das. Slotted symmetric power control in managed wlans. *Technical report, NEC Laboratories America.* 6, 6.1.1

[87] Vishnu Navda, Anand Prabhu Subramanian, Kannan Dhanasekaran, Andreas Timm-Giel, and Samir Das. MobiSteer: using steerable beam directional antenna for vehicular network access. In *MobiSys*, 2007. 1.1.3, 2.2.1, 4.5.1, 4.5

[88] Jeffrey Pang, Ben Greenstein, Damon McCoy, Michael Kaminsky, and Srinivasan Seshan. Wifi-reports: Improving wireless network selection with collaboration. In *MobiSys*, 2009. 5.3.4

[89] K. Papagiannaki, M. Yarvis, and W. S. Conner. Experimental characterization of home wireless networks and design implications. In *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, pages 1 –13, 2006. doi: 10.1109/INFOCOM.2006.293. 2.1, 2.6

[90] Rabin K. Patra, Sergiu Nedevschi, Sonesh Surana, Anmol Sheth, Lakshminarayanan Subramanian, and Eric A. Brewer. WiLDNet: Design and implementation of high perfor-

mance WiFi based long distance networks. In *NSDI*, 2007. 1.1.1, 1.1.2, 1.1.3, 2.2.1, 2.2.3, 4, 4.3.2, 4.5.1, 4.5

[91] A. Prabhu and S.R. Das. Addressing deafness and hidden terninal problem in directional antenna based wireless multi-hop networks. *COMSWARE*, Jan. 2007. 4.5.1

[92] Daji Qiao, Sunghyun Choi, Amit Jain, and Kang G. Shin. Miser: an optimal low-energy transmission strategy for ieee 802.11a/h. In *MobiCom*, 2003. 6, 6.1.1

[93] Bozidar Radunovic and Jean-Yves Le Boudec. Rate performance objectives of multihop wireless networks. *IEEE Transactions on Mobile Computing*, 3:334–349, October 2004. ISSN 1536-1233. 1.1.3, 2.1, 4.5.4

[94] Ramya Raghavendra, Elizabeth M. Belding, Konstantina Papagiannaki, and Kevin C. Almeroth. Understanding handoffs in large ieee 802.11 wireless networks. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, IMC '07, pages 333–338, 2007. 5.2.2

[95] Kishore Ramachandran, Ravi Kokku, Honghai Zhang, and Marco Gruteser. Symphony: synchronous two-phase rate and power control in 802.11 wlans. In *MobiSys*, 2008. 1.1.1, 1.1.3, 1.1.3, 3.3, 6, 6, 6.1.1, 6.1.2, 6.2.1, 2

[96] Kishore Ramachandran, Ravi Kokku, Karthikeyan Sundaresan, Marco Gruteser, and Sampath Rangarajan. R2d2: regulating beam shape and rate as directionality meets diversity. In *MobiSys '09: Proceedings of the 7th international conference on Mobile systems, applications, and services*, pages 235–248. ACM, 2009. 2.2.1

[97] Bhaskaran Raman and Kameswari Chebrolu. Design and evaluation of a new MAC protocol for long-distance 802.11 mesh networks. In *MobiCom*, 2005. 1.1.2, 2.2.1, 2.2.3, 4, 4.5.1

[98] Aniruddha Rangnekar and Krishna M. Sivalingam. Qos aware multi-channel scheduling for ieee 802.15.3 networks. *Mob. Netw. Appl.*, 11:47–62, February 2006. 4.5.4

[99] Ashish Raniwala and Tzi-Chiueh. Architecture and algorithms for an ieee 802.11-based multi-channel wireless mesh network. In *Infocom*, 2005. 5.2.2

[100] Theodoros Salonidis and Leandros Tassiulas. Distributed dynamic scheduling for end-to-end rate guarantees in wireless ad hoc networks. In *Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '05, pages 145–156, 2005. 4.5.4

[101] Theodoros Salonidis, Leandros Tassiulas, and Ros Tassiulas. Distributed on-line schedule adaptation for balanced slot allocation in wireless ad hoc networks. Technical report, In Proc. IEEE International Workshop on Quality of Service (IWQoS), Montreal,Canada, 2002. 4.5.4

[102] Aimin Sang, Xiaodong Wang, Mohammad Madihian, and Richard D. Gitlin. Coordinated load balancing, handoff/cell-site selection, and scheduling in multi-cell packet data systems. In *Proceedings of the 10th annual international conference on Mobile computing and networking*, MobiCom '04, pages 302–314, 2004. 5.2.2

[103] Naveen Kumar Santhapuri, Justin Manweiler, Souvik Sen, Romit Roy Choudhury, Srihari

Nelakuditi, and Kamesh Munagala. Message in Message MIM: A case for reordering transmissions in wireless networks. In *HotNets*, 2008. 3.2.1, 3.3

[104] Vasudev Shah and Srikanth Krishnamurthy. Handling asymmetry in power heterogeneous ad hoc networks: A cross layer approach. In *ICDCS*, 2005. 6, 6.1.1

[105] Gaurav Sharma, Ravi R. Mazumdar, and Ness B. Shroff. On the complexity of scheduling in wireless networks. In *Proceedings of the 12th annual international conference on Mobile computing and networking*, MobiCom '06, pages 227–238, 2006. 4.2.4

[106] A. Sheth and R. Han. Shush: reactive transmit power control for wireless mac protocols. *WICON*, July 2005. 6, 6.1.1

[107] Anmol Sheth and Richard Han. Adaptive power control and selective radio activation for low-power infrastructure-mode 802.11 lans. In *ICDCSW*, 2003. 6, 6.1.1

[108] Anmol Sheth, Sergiu Nedevschi, Rabin K. Patra, Sonesh Surana, Eric A. Brewer, and Lakshminarayanan Subramanian. Packet loss characterization in WiFi-based long distance networks. In *INFOCOM*, 2007. 2.2.1, 2.4.3, 4

[109] Vivek Shrivastava, Nabeel Ahmed, Shravan Rayanchu, Suman Banerjee, Srinivasan Keshav, Konstantina Papagiannaki, and Arunesh Mishra. Centaur: realizing the full potential of centralized wlans through a hybrid data path. In *MobiCom '09: Proceedings of the 15th annual international conference on Mobile computing and networking*, pages 297–308, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-702-8. doi: http://doi.acm.org/10.1145/1614320.1614353. 3.3, 5.3.2, 5.3.5, 3, 6.4.2

[110] Pablo Soldati, Björn Johansson, and Mikael Johansson. Proportionally fair allocation of end-to-end bandwidth in stdma wireless networks. In *Proceedings of the 7th ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '06, pages 286–297, 2006. ISBN 1-59593-368-9. 1.1.3, 4.5.4

[111] Ion Stoica, Scott Shenker, and Hui Zhang. Core-stateless fair queueing: a scalable architecture to approximate fair bandwidth allocations in high-speed networks. *IEEE/ACM Trans. Netw.*, 11:33–46, February 2003. 2.1

[112] Anand Prabhu Subramanian, Pralhad Deshpande, Jie Gao, and Samir R. Das. Drive-by localization of roadside WiFi networks. In *INFOCOM*, May 2008. 2.2.1, 4.5.3

[113] Anand Prabhu Subramanian, Henrik Lundgren, and Theodoros Salonidis. Experimental characterization of sectorized antennas in dense 802.11 wireless mesh networks. In *MobiHoc*, 2009. 8.2.3

[114] Karthikeyan Sundaresan and Konstantina Papagiannaki. The need for cross-layer information in access point selection algorithms. In *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*, IMC '06, pages 257–262, 2006. 5.2.2

[115] Karthikeyan Sundaresan and Konstantina Papagiannaki. The need for cross-layer information in access point selection algorithms. In *IMC*, 2006. 5.3.4

[116] Karthikeyan Sundaresan and Raghupathy Sivakumar. A unified MAC layer framework for ad-hoc networks with smart antennas. *IEEE/ACM Trans. Netw.*, 15(3), 2007. 3.1, 3.3

[117] Karthikeyan Sundaresan, Weizhao Wang, and Stephan Eidenbenz. Algorithmic aspects of

communication in ad-hoc networks with smart antennas. In *MobiHoc*, 2006. 4.5.2

[118] Karthikeyan Sundaresan, Kishore Ramachandran, and Sampath Rangarajan. Optimal beam scheduling for multicasting in wireless networks. In *MobiCom '09: Proceedings of the 15th annual international conference on Mobile computing and networking*, pages 205–216, New York, NY, USA, 2009. ACM. 2.2.1, 4.5.2

[119] Mineo Takai, Jay Martin, Rajive Bagrodia, and Aifeng Ren. Directional virtual carrier sensing for directional antennas in mobile ad hoc networks. In *MobiHoc*, 2002. 4.5.1

[120] S. Vasudevan, K. Papagiannaki, C. Diot, J. Kurose, and D. Towsley. Facilitating access point selection in ieee 802.11 wireless networks. In *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*, IMC '05, pages 26–26, 2005. 5.2.2

[121] Mythili Vutukuru, Kyle Jamieson, and Hari Balakrishnan. Harnessing exposed terminals in wireless networks. In *NSDI*, 2008. 3.1, 3.3, 6.3.2

[122] J. Wang, Y. Fang, and D. Wu. SYN-DMAC: a directional MAC protocol for ad hoc networks with synchronization. *MILCOM 2005*. 4.5.1

[123] K. Whitehouse, A. Woo, F. Jiang, J. Polastre, and D. Culler. Exploiting the capture effect for collision detection and recovery. In *Proceedings of the IEEE EmNetS-II*, May 2005. 3.2.1

[124] Mehmet Yavuz, Farhad Meshkati, Sanjiv Nanda, Akhilesh Pokhariyal, Nick Johnson, Balaji Raghothaman, and Andy Richardson. Interference management and performance analysis of umts/hspa+ femtocells. *Comm. Mag.*, 47:102–109, September 2009. 6.1.3

[125] Ji-Hoon Yun and Kang G. Shin. Ctrl: a self-organizing femtocell management architecture for co-channel deployment. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, MobiCom '10, pages 61–72, 2010. 6.1.3

[126] Yan Zhang, Honglin Hu, and Hsiao-Hwa Chen. Qos differentiation for ieee 802.16 wimax mesh networking. *Mob. Netw. Appl.*, 13:19–37, April 2008. 4.5.4

[127] Zhensheng Zhang. DTRA: directional transmission and reception algorithms in WLANs with directional antennas for QoS support. *Network, IEEE*, 19(3), May-June 2005. 4.5.1

[128] R. Zheng and R. Kravets. On-demand power management for ad hoc networks. *INFOCOM*, 2003. 6, 6.1.1

[129] Gang Zhou, Tian He, Sudha Krishnamurthy, and John A. Stankovic. Impact of radio irregularity on wireless sensor networks. In *MobiSys '04*, pages 125–138. ACM, 2004. ISBN 1-58113-793-1. doi: http://doi.acm.org/10.1145/990064.990081. 2.1, 2.6

[130] Jun Zou and Dongmei Zhao. Bottleneck-first scheduling for real-time traffic in ieee 802.11 infrastructure-based mesh networks. In *Proceedings of the 2006 international conference on Wireless communications and mobile computing*, IWCMC '06, pages 593–598, 2006. 4.5.4

[131] Jun Zou and Dongmei Zhao. Real-time voice traffic scheduling and its optimization in ieee 802.11 infrastructure-based wireless mesh networks. In *Proceedings of the 3rd international conference on Quality of service in heterogeneous wired/wireless networks*, QShine '06, 2006. 4.5.4